



12-2012

## From Building Evaluation Capacity to Supporting Evaluation Capacity Development: The Cases of Democratic Republic of Congo, Niger, and South Africa

Michele Tarsilla  
Western Michigan University, [michele\\_tarsilla@yahoo.com](mailto:michele_tarsilla@yahoo.com)

Follow this and additional works at: <https://scholarworks.wmich.edu/dissertations>



Part of the African Studies Commons, and the Policy Design, Analysis, and Evaluation Commons

---

### Recommended Citation

Tarsilla, Michele, "From Building Evaluation Capacity to Supporting Evaluation Capacity Development: The Cases of Democratic Republic of Congo, Niger, and South Africa" (2012). *Dissertations*. 122.  
<https://scholarworks.wmich.edu/dissertations/122>

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact [wmu-scholarworks@wmich.edu](mailto:wmu-scholarworks@wmich.edu).



FROM BUILDING CAPACITY TO SUPPORTING EVALUATION CAPACITY  
DEVELOPMENT: THE CASES OF DEMOCRATIC REPUBLIC OF CONGO, NIGER  
AND SOUTH AFRICA

by

Michele Tarsilla

A Dissertation  
Submitted to the  
Faculty of The Graduate College  
in partial fulfillment of the  
requirements for the  
Degree of Doctor of Philosophy  
Interdisciplinary Ph.D.in Evaluation  
Advisor: Chris Coryn, Ph.D.

Western Michigan University  
Kalamazoo, Michigan  
December 2012

FROM BUILDING EVALUATION CAPACITY TO SUPPORTING EVALUATION  
CAPACITY DEVELOPMENT: THE CASES OF DEMOCRATIC REPUBLIC  
OF CONGO, NIGER AND SOUTH AFRICA

Michele Tarsilla, Ph.D.

Western Michigan University, 2012

Building on both the current thinking among international development practitioners and the on-going scholarly debate on evaluation capacity development (ECD), this study aims to identify strategies that might help to strengthen national evaluation capacity in a variety of countries in a more inclusive and sustainable fashion in the future. Based on a case study design featuring an extensive literature review of specialized literature (both within and outside the evaluation field), a series of semi-structured interviews and three rounds of on-line validation sessions held with ECD researchers, funders and implementers; this study aims at three main objectives. First, in an attempt to enhance a more proper use of the term ECD as opposed to that of evaluation capacity building (ECB), several ECD central attributes are identified. Second, based on data collection carried out in Niger, Democratic Republic of Congo and South Africa, the capabilities (and corresponding weaknesses) of national evaluation associations as well as the magnitude of their contribution to national ECD programming, are assessed. Third, based on the analysis of the shortcomings associated with some of the existing ECD evaluative frameworks, some essential criteria for measuring ECD results more effectively in the future are presented. Results show that, for the sake of promoting a national evaluative culture in international development contexts

more successfully in the future, two main strategies ought to be pursued. First, to make ECD programming more participatory, a plurality of governmental and non-governmental actors need to be involved in both the design and implementation of ECD programs. Second, in order to enhance info-sharing and evaluation knowledge-building among a variety of ECD stakeholders both in the public and private sector, the feasibility of implementing sphere-crossing initiatives within the scope of national ECD program needs to be explored more systematically. Such is the case of activities and programs supporting the creation and/or strengthening of formal and informal national evaluation associations, increasingly referred to by development agencies as Voluntary Organization of Professionals in Evaluations (VOPE).

Copyright by  
Michele Tarsilla  
2012

## ACKNOWLEDGMENTS

A project of the magnitude of a doctoral dissertation is not a solitary endeavor and is possible only with the help of many people. Above all, I am deeply indebted to my doctoral committee: Dr. Chris Coryn, (the Committee Chair), Dr. Michael Bamberger and Mr. Jim Rugh. Thank you.

I will be eternally grateful to Dr. Coryn (the Chair of my doctoral committee) for the support and continued encouragement he has provided me with since I entered the Interdisciplinary Doctoral Program in Evaluation (IDPE) at Western Michigan. As a dedicated professor and pragmatic thinker, Dr. Coryn pushed me constantly to think about evaluation methods and processes outside of what used to be my comfort zone (the application of mixed methods to evaluations of international development projects).

Equally important was the contribution of Dr. Bamberger with whom I was fortunate to work on a number of evaluation assignments in Moldova and El Salvador. Dr. Bamberger embodied all the traits that I have always thought a brilliant evaluator should possess: a quick mind, a robust methodological grounding, an innate curiosity for local cultures and programs' intangible aspects and, above all, great humility.

Just as crucial was Mr. Rugh's contribution to the successful pursuit of my doctoral research. He and the work he had performed with national evaluation associations around the world over in recent decades inspired me to focus on the value and practical relevance of sound evaluation capacity development programs in international development contexts. It was through my discussions with him that I developed the idea of conducting three case studies on Voluntary Organizations Promoting Evaluation (VOPE) in sub-Saharan Africa.

## Acknowledgments—Continued

Overall, the time invested by my committee in reviewing and offering suggestions for my manuscript under quite tight deadlines made this dissertation research and the whole writing process a truly eye-opening, educational and fulfilling experience. Once again, I must thank my Committee. I am fortunate that I had the opportunity to benefit from engaging exchanges with some of the world's top scholars and practitioners in the international development field during my research. I am especially grateful to Caroline Heider (currently one of the World Bank Vice Presidents and the Director of the Independent Evaluation Group) and Zenda Ofir (former President of the African Evaluation Association and one of the founders of the South African National Evaluation Association). In particular, I appreciate their time and interest in providing me with feedback on some of the emerging themes in my study despite their hectic work and travel schedule.

Other international evaluation capacity development specialists who influenced my thinking through their work and their communications with me, and whom I am very grateful to, are Richard Blue (Senior Vice President at Social Impact Inc.); Fred Carden (International Development Research Center, Canada); Ian Hopwood (ex-Unicef Country Director and currently one of the strongest supporters of national evaluation associations in Africa); Colin Kirk (current Director of the UNICEF Evaluation Office); Keith McKay (former World Bank Lead Evaluation Capacity Development Specialist); Penny Hawkins, (Deputy Direct of the Evaluation Office at the Rockefeller Foundation); Mohamed Manai (Manager at the African Development Bank Development Office); Indran Naidoo (Director of the UNDP Evaluation Office); Stephen Porter (CLEAR Initiative South Africa); Sukai Prom-Jackson (Advisor at the UNDP Evaluation Office); Jean-Serge Quesnel (former Director of the UNICEF Evaluation Office); Romeo Santos (WorkLand Institute, the Philippines) Belen Sanz Luque (Chief of the UNWomen Evaluation Office); Rolf Sartorius

## Acknowledgments—Continued

(President at Social Impact Inc.) and Nicoletta Stame (former President of the European Evaluation Society).

A special thanks also goes to the colleagues from the three VOPE that conducted my case studies on. In particular, I would like to acknowledge Albert Mashika and Firmin Seke in the Democratic Republic of Congo; Gado Boureima, Chefou Balla Idrissa and Yayé Seidou in Niger; Terency Beney, Jennifer Bisgard and Thabo Mabogoane in South Africa.

A special thank goes to all the members of the International and Cross-Cultural Evaluation Topical Interest Group at the American Evaluation Association who supported me throughout my research and who shared with me some of the best “critical friends” perspectives that a doctoral candidate may ever wish for. In particular, the feedback provided on some of my findings and conclusions by Oumoul Khayri Ba Tall and Florence Etta (both former Presidents of the African Evaluation Association) as well as Craig and Karen Russon (Senior Evaluation Officer at the International Labor Organization and President of the Evaluation Capacity Development Group respectively) were invaluable.

Next, I would not have been able to get to this point if I had not met some six great mentors before entering the program, whom I would also like to acknowledge. First, Linda Morra and Ray Rist, my World Bank colleagues who strongly encouraged me to first attend IPDET in Ottawa and then to pursue a doctoral degree in evaluation. Second, Rosalia Rodriguez-Garcia, whose intellectual rigor and leadership during my work with her at the World Bank on the evaluation of the community response on HIV and AIDS were a true inspiration in my professional career path. Third, Bill Roche, whose willingness to be there to listen and to help regardless of the issues at stake (including editing one of my chapters on a very tight deadline) confirmed to me that kindness and altruism still exist. Fourth, Ed Avant, whose continued friendship and affection represented my “home away from home”



## Acknowledgments—Continued

in DC, despite the numerous international missions I undertook over the last ten years.

Fifth, Laura Nyblade, whose friendship together with her professional and dissertation- related advice meant a great deal to me this past year.

Furthermore, I would also like to extend my thanks to some of the staff associated with the Interdisciplinary Doctoral Program in Evaluation (IDPE) at Western Michigan University: Dr. Daniela Schröter (Director of Research at IDPE) for making my life in the Midwest an enjoyable experience and Mary Ramlow (IPDE Admission Officers) for turning the overall dissertation writing process into a more manageable experience. Likewise, I owe much to my dear friends Therese Fergo, Josh Hughes, Trevor Lockamy, Kelly Robertson and Diane Rogers for their continued support over the past few years.

In conclusion, I offer my deepest gratitude to my parents, Mario and Cecilia Tarsilla, who taught me that hard work coupled with integrity and a sense of adventure could help you get anywhere you want in life. Grazie per sempre papà e mamma!

Michele Tarsilla

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	ii
LIST OF TABLES.....	xi
LIST OF FIGURES .....	xii
 CHAPTER	
1. INTRODUCTION.....	1
Statement of the Problem.....	4
Aim and Scope of the Dissertation.....	12
Organization of the Dissertation.....	15
 2. REVIEW OF THE LITERATURE .....	 17
Introduction.....	18
The Current State of ECB and ECD Literature .....	19
The ECB and ECD Definitional Dilemma .....	22
Definitions of Terminology .....	28
The Inclusiveness of ECB and ECD Targeting International Development Contexts.....	29
Introduction.....	29
General Inquiry Framework.....	30
The Biased Targeting of ECD and ECB Programming.....	31
VOPE’s Involvement in ECD Programming.....	35
Evaluation’s Role in Addressing the Biased Targeting of ECD Programming .....	40
Main ECB and ECD Frameworks.....	41
The ECB Supply-Demand Equation and the Evaluation Capacity Development Framework of Action.....	42
The Multidisciplinary ECB Model.....	47
The Three-level ECD Framework.....	51
The Vygotskian ECB Model .....	54
The Need for a New ECB/ECD Framework .....	55
ECD Evaluation Frameworks .....	56
Introduction.....	56
Evaluative Frameworks and Approaches .....	56
 3. METHODS.....	 70
Design .....	71
Data Collection Methods.....	74
Sample.....	74
Procedure .....	75
Data Processing and Analysis.....	76

## Table of Contents—Continued

CHAPTER		
	Semi-structured Interviews.....	78
	Sample.....	80
	Instrumentation.....	82
	Procedure.....	83
	Analysis.....	84
	Use of Triangulation and Scenario of Convergence/Divergence.....	84
	Semi-structured Interviews with VOPE Coordinators and Other VOPE Members.....	85
	Sample.....	85
	Instrumentation.....	88
	Procedure.....	89
	Analysis.....	89
	Review of VOPE Documentation.....	91
	Sample.....	91
	Instrumentation.....	92
	Procedure.....	92
	Analysis.....	92
4.	FINDINGS FROM THE SYSTEMATIC LITERATURE REVIEW.....	93
	The Relationship between ECB and ECD.....	93
	Findings from the Systematic Literature Review.....	93
	Capacity.....	94
	Capability.....	102
	Capacity Development.....	106
	Evaluation Capacity Building.....	110
	Evaluation Capacity Development.....	117
	Conclusions of Literature Review on ECB and ECD Definitions.....	121
5.	VOPE'S ROLE IN NATIONAL ECD PROGRAMS.....	125
	Introduction.....	125
	Case Study 1: South Africa Monitoring and Evaluation Association (SAMEA).....	126
	Introduction.....	126
	SAMEA History.....	127
	SAMEA Capability Assessment.....	133
	Other Stakeholders in SA and Corresponding Institutional Arrangement.....	139
	THE DPME.....	139
	Chapter Nine Institutions.....	141
	The SA Public Administration Leadership and Management Academy (PALAMA).....	144
	The M&E Learning Network.....	147
	ECD Stakeholders Outside the Government Spheres.....	149
	Relevant Contextual Factors Contributing to the Development of a Strong	

## Table of Contents—Continued

CHAPTER		155
	Evaluative Culture/Enhancing ECD in SA.....	155
	Factors Discouraging the Development of a Strong Evaluative Culture/ECD in SA.....	157
6.	CASE STUDY 2: NIGER MONITORING AND EVALUATION NETWORK (ReNSE).....	159
	Introduction.....	159
	ReNSE History.....	159
	ReNSE Capability Assessment.....	165
	Other ECD Stakeholders in Niger.....	173
	Capacity Building Providers in Niger I: International Development Partners.....	173
	Capacity Building Providers in Niger II: National Training Institutions.....	177
	Plans for the Future.....	182
	Capacity Building Providers III: Research and Planning Divisions within Ministries.....	185
	Relevant Contextual Factors Contributing to the Development of a Strong Evaluative Culture/Enhancing ECD in Niger.....	185
	The Establishment of Participatory Planning Processes.....	185
	Factors Discouraging the Development of a Strong Evaluative Culture/ECD in Niger.....	191
	ReNSE Membership: Do Ownership and Inclusiveness Always Go Together?.....	196
	Pros and Cons of Formalizing an Informal Network: The ReNSE Experience.....	197
	The Future of ECD in Niger.....	201
7.	CASE STUDY 3: THE DEMOCRATIC REPUBLIC OF CONGO MONITORING AND EVALUATION ASSOCIATION (ACoSE).....	204
	Introduction.....	204
	ACoSE History.....	206
	ACoSE Capability Assessment.....	210
	Other ECD Stakeholders in the DRC.....	214
	Factors Discouraging the Development of an Evaluative Culture/ECD in the DRC.....	221
8.	CURRENT PRACTICES TO MEASURE ECD.....	229
	Introduction.....	229
	Variables and Levels.....	230
	The First Set of ECD Evaluative Variables.....	232
	The Second Set of ECD Evaluative Variables.....	233
	The Third Set of ECD Evaluative Variables.....	236

## Table of Contents—Continued

CHAPTER	
The Fourth Set of ECD Evaluative Variables.....	239
Are All ECD Effects Measurable? .....	240
Difficult Comparability across ECD Interventions .....	243
VOPEs Evaluation .....	243
First VOPEs Outcome: Knowledge Building .....	245
Second VOPEs Objective: Professionalization of In-country Evaluation Practices.....	245
Third VOPEs Objective: Enhancing Institutional Relations .....	246
Measuring VOPEs Organizational Development.....	247
9. CONCLUSIONS .....	248
First Research Question .....	248
Going Beyond the Existing Paradigms .....	253
Second Research Question.....	258
Organizational Processes .....	258
Organizational Capabilities .....	260
Current ECD Frameworks’ Limitations and Opportunities for More Inclusive Targeting in the Future.....	267
Good Practices to Enhance ECD Inclusiveness .....	268
Third Research Question.....	271
How Can, or How Should, ECD Best Be Evaluated?.....	271
Five Key Principles in Evaluating ECD Programming.....	274
Study Limitations .....	277
Implications and Theory and Practice.....	279
Future Research .....	280
APPENDICES	
A. Acronyms .....	282
B. Semi-structured Interview Protocol .....	286
C. The Five Key Dimensions of the ECPDM Capability Assessment.....	288
D. VOPE Capability Assessment Tool .....	291
E. HSIRB Approval Letter.....	294
F. ECD Continuum Feedback Questionnaire.....	296
G. Review of Literature on Capacity: Summary Key Findings .....	298
H. Review of Literature on Capability: Summary Key Findings .....	301
I. Review of Literature on Capacity Development: Summary Key Findings.....	303
J. The Evaluation Capacity Development Continuum Framework (ECD-CF).....	305
K. The SFAR Framework and the ECD Ecology in the DRC .....	315
L. The SFAR Framework and the ECD Ecology in Niger .....	317

Table of Contents—Continued

APPENDICES	
M. The SFAR Framework and the ECD Ecology in South Africa.....	319
REFERENCES.....	321

## LIST OF TABLES

1.	Limitations in Measuring CD Internationally: The Case of the Paris Declaration on Aid Effectiveness.....	11
2.	ECB And ECD Definition And Hermeneutical Gaps .....	27
3.	The ECB Equation.....	43
4.	The Four Forces Influencing ECB Capacity (By Level And Role).....	49
5.	Kirkpatrick Four Level Evaluation Model.....	57
6.	ECB/ECD Matrix.....	78
7.	Respondents' Profile .....	81
8.	VOPE Organizational Processes: Comparative Checklist .....	89
9.	VOPE Main Activity Main Objectives.....	90
10.	SAMEA Membership (By Category).....	131
11.	Consistency Between SAMEA Mission/Objectives and Activities.....	134
12.	The ISPE Curriculum: Course and Corresponding Credits.....	181
13.	Kirkpatrick 4-Level Model: Level 1 and Level 2.....	233
14.	Kirkpatrick 4-Level Model: Level 3 and 4.....	234
15.	VOPEs Organizational Processes: A Comparative Checklist .....	259
16.	VOPEs Organizational Capabilities: A Comparative Summary.....	260
17.	VOPEs Main Objectives and Outcomes: A Comparative Summary .....	273

## LIST OF FIGURES

1.	A Multidisciplinary Model of Evaluation Capacity Building (ECB) .....	47
2.	The Three Level Model: Key Elements And Major Contributing Authors .....	52
3.	ECB Results Framework .....	60
4.	Six Intermediate Capacity Outcomes .....	61
5.	Examples of Predictive Capacity In ECB Planning .....	62
6.	Sequencing of Qualitative Methods.....	79
7.	The SFAR Framework.....	225
8.	The Four Sets of ECB/ECD Evaluative Variables.....	231
9.	Apparent Lack of Logical Causal Pathways.....	244
10.	Theory of Change Behind First VOPEs Outcome .....	245
11.	Theory of Change Behind Second VOPEs Outcome .....	246
12.	Theory of Change Behind Third VOPEs Outcome .....	247
13.	A Simple ECB Programming.....	249
14.	A More Comprehensive ECB Programming .....	250
15.	A True ECD Programming.....	251



## CHAPTER 1

### INTRODUCTION

Evaluation policies and strategic evaluation plans currently in use among several development agencies around the world are predicated on the assumption that international development evaluation serves two primary functions (German Development Cooperation, 2012; Norwegian Agency for Development Cooperation, 2006; Swedish International Development Cooperation Agency 2007; United States Agency for International Development 2011). First, to enhance the accountability of those who manage and implement international development projects, especially vis-à-vis their respective funders<sup>1</sup> and expected beneficiaries (Wiesner, 1997). Second, to foster learning (among those who commission, manage, conduct, and use evaluation) on what works well and what needs to be improved in international development projects and programs (Argyris et al., 1996; Bamberger, 2009, Pasteur, 2006; Solomon & Chowdhury, 2002).

Based on such assumptions, any activity aimed at strengthening evaluation function<sup>2</sup>—locally, nationally, or globally—should, therefore, be able to contribute to

---

<sup>1</sup> The term funders used in this study includes a) direct funders, such as Non Governmental Organizations (NGOs) or in-country agencies serving as an intermediary of larger organizations for disbursement purposes; and b) indirect funders, such as citizens whose taxes or direct donations to NGOs and other types of development organizations are being used to fund development intervention.

<sup>2</sup> A good illustration of a strong national evaluation capacity is the combination of a strong government demand for evaluation, a systematic use of evaluation findings, a sufficiently large supply of in-country evaluation experts providing their professional services to meet the in-country demand for evaluation, and the building of local evaluation knowledge through national evaluation associations.

strengthening both the performance and effectiveness of international development projects<sup>3</sup> in a variety of countries. However, this is easier said than done. Typically, funders and international development agencies attempt to strength the evaluation function by developing the technical skills of international development practitioners. However, such strategies do not always translate into stronger development effectiveness. One reason for this is the lack of a genuine evaluative culture (e.g., the systematic conduct of evaluation and the use of findings for decision-making), often resulting from the limited ability of Evaluation Capacity Building (ECB) to foster ownership and inclusiveness of evaluation processes (Trochim, 1991).

The scenario, however, is not as bleak as would first appear. There are several examples of countries that have put in place Monitoring and Evaluation (M&E) systems that are both prospering as well as serving accountability and learning purposes. Chile, Columbia, Malaysia, Mexico, Sri Lanka (Wijayatilake, 2011), and South Africa are good illustrations of how the creation of a supportive environment to enhance the evaluation function at the organizational (Stevenson, 2002) and institutional levels can foster the development of a strong national “evaluative culture” (Boyle & Lemaire, 1999; Mayne, 2008; Trochim, 2006).<sup>4</sup> One feature that all these countries have in common is the buy-in of different stakeholder groups (both within and outside the national government) into the evaluation discourse and

---

<sup>3</sup> That is especially relevant at a time when international development resources are decreasing. Breaking a long trend of annual increases, members of the Organization for Economic Co-operation and Development European (OECD) Development Assistance Committee (DAC) allocated US\$133.5 billion of net official development assistance (ODA) in 2011, that is a 2.7% drop in real terms compared to 2010 (OECD, 2011a).

<sup>4</sup> The experiences of setting up and implementing a variety of multi-level and participatory evaluation policies (Atkinson et al., 2005) and processes in Bogota, Colombo, Kuala Lumpur, Mexico City, Pretoria, and Santiago; are certainly limited in scope and do not lend themselves to warranting conclusions and recommendations directly applicable to other countries. However, one recurrent idea that emerged in all these six countries and that might want to be explored further for future applicability in other contexts, is the correlation existing between the proliferation of an evaluative culture and the degree of involvement in it of citizens - either individually or organized in Civil Society Organizations (CSO) – as well as local and national institutions (Burton, 2009; Reeler, 2007, Rodriguez-Biella & Monterde-Diaz, 2010).

their subsequent involvement in all the related processes. That notwithstanding, such success stories have not been capitalized on in the international development arena as frequently as they should, of could, have. In an effort to understand the origins of this, the rigidity of international organizations' missions and structures has been identified as one of the main barriers to the inclusiveness of Evaluation Capacity Development (ECD) processes; that is, to the involvement of actors operating both within and outside national governments. On the one hand, cognizant of the apparent link existing between the participatory nature of national evaluation processes and the strength of in-country evaluation culture, bilateral donors and philanthropic foundations have been able to fund<sup>5</sup> a plethora of initiatives and programs specifically aimed at creating a more favorable environment for the strengthening of evaluation knowledge and skills among a variety of stakeholders (including representatives from academia, the private sector and non-governmental organizations) for over a decade (OECD 2006a).<sup>6</sup> On the other hand, multilateral agencies (e.g., United Nations, the World

---

<sup>5</sup> Although accurate figures on the total amount that donors spend on Evaluation Capacity Development (ECD) are not available, the rather conservative estimate proposed in this study (\$500 million per year or a little bit over 3% of total annual donor aid that is specifically allocated for technical cooperation) and based on the data provided by the European Capacity Development Program Management (ECDPM) (Baser & Morgan, 2008) and OECD, (OECD, 2006b)<sup>5</sup>, attests to the magnitude of both the scope and relevance of ECD in international development today<sup>5</sup>.

<sup>6</sup> Despite the difficulties encountered by many development agencies in quantifying the total amount of budgeted resources and costs associated with the implementation of ECD, the 2011 figures seem to confirm a positive trend in the level of funding allocated by international donors to enhancing overall capacity in developing countries since the early 2000s. According to the estimates released by the Organization for Security and Co-operation in Europe (OECD) Development Assistance Committee (DAC), the total amount of resources allocated by DAC members for technical co-operation in developing countries amounted to US\$18.4 billion in 2003 (27% of the total net budget Official Development Assistance or ODA) and US\$20.8 billion in 2004 (27% of total net budget ODA) (OECD, 2012). Furthermore, the pivotal role played by capacity development interventions within the portfolio of several international donors has been confirmed for years by a variety of research publication either commissioned by bilateral agencies themselves or multilateral organizations. In the case of the Canadian International Development Agency (CIDA), a review of the agency's portfolio (Lavergne, Lewis et al., 2004) stated that as much as 74% of all its activities (weighted by disbursements) could be classified as promoting capacity development. Similarly, according to the same study, 76% of all the projects funded by the International Development Research Center (IDRC) included some capacity development component. In the case of the multilateral organizations, the cases of the Asian Development Bank (ADB) and the World Bank (WB) are quite illustrative. According to a report published by the United National Development Programme (UNDP, 2006), as of 2005, the ADB Capacity development

Bank) have allocated the largest share of ECD resources to national governments rather than civil society because of their specific political and organizational mandate.

Far from becoming rhetoric or a purely academic exercise,<sup>7</sup> an exploration of ECD central attributes, modalities of implementation and evaluative variables, would be particularly useful to enhance ECD programming in the future. This is especially relevant given that capacity development, in evaluation as well as in a variety of other fields, is not the only “missing link in development,” (World Bank, 2005, p. 24) but even more importantly, part of the overall goal of development cooperation (Fukuyama, 2004).<sup>8</sup>

### Statement of the Problem

Funding made available by development partners and national governments for the implementation of ECD programs within the scope of international development contexts has increased over the last decade (OECD, 2011b). However, the effectiveness of activities aimed at supporting national evaluation capacity has been hindered by three main factors.

The first hindrance is the lack of a common, agreed-upon definition of ECD goals and implementation modalities among both national governments and development partners. Such definitional and epistemological ambiguity has been exacerbated by the

---

budget amounted to about one fifth of total lending for Technical Assistance (TA) (almost US\$40 million), which constituted approximately 13% of its total lending (Qureshi et al., 2008). In the case of the World Bank, a report assessing the effectiveness of training in 179 projects funded by the agency between 2002 and 2006, (World Bank Independent Evaluation Group, 2008) estimated that six percent of the agency’s overall lending budget was allocated to training (US\$155 million per year).

<sup>7</sup> Some practitioners, especially in the US, might criticize any debate or inquiry focused on the development of a new ECD definition as unnecessary question of semantic finesse. However, several field practitioners who had been interviewed before the start of the study expressed their disagreement with the use of ECB and ECD as synonyms and said that there was a need for a clearer definition of what ECD is and how it works.

<sup>8</sup> The increased interest in evaluation capacity development was also accompanied by a surge in debates on evaluation capacity held within the evaluation community. Since 2000, over 200 presentations on evaluation capacity building (ECB) have been given at the conferences of the American Evaluation Association (AEA) and Canadian Evaluation Society (CES) (Boyle, 2005; Preskill, 2008; Preskill & Boyle, 2008). Interestingly, 73 sessions and posters on capacity building (out of a total of 1,000) were presented during the 2012 AEA Annual conference. Similarly, membership within the Learning Organization and Capacity Building Topical Interest Group (TIG) within AEA, grew exponentially over the last few years (with its 1,000 members, the TIG was the third largest within AEA as of 2012).

paucity of peer-reviewed literature on ECD and has been further aggravated by the practice of many evaluation practitioners (especially in the United States) who use the term ECD as a synonym for evaluation capacity building (ECB) or evaluation skills development.<sup>9</sup> As a result, despite many international development evaluators asserting that ECD is more comprehensive and contextually relevant to national development process than ECB, the majority of donor-funded ECD activities implemented in the field have been assimilated with either workshops or coaching sessions aimed at government officials (Bemelmans-Videc et al., 2003; Cracknell, 2000; Morgan, 2006). As a result, the systemic nature of ECD has not been fully understood by a variety of actors in international development and donors' ECD efforts have not been harmonized (Easterley, 2007).<sup>10</sup>

In response to such weakness in ECD programming, there appears to be the need for a new definition of ECD<sup>11</sup> that could fill the gap between the scant peer reviewed literature available on ECD and the more copious grey literature available on the topic,<sup>12</sup> as well as between the looser connotation assigned to the term ECD in the U.S. and its more ideological characterization in the rest of the world (Liverani & Lundgren, 2007). The relevance of such need is all the more apparent as a similar gap in the area of ECD was

---

<sup>9</sup> This is especially true in the United States where the term ECD, often perceived to be a synonym of Evaluation Capacity Building (ECB), capacity strengthening or, national monitoring and evaluation system building, has never been used as widely as in the rest of the world (Blue et al., 2009a; Blue et al., 2009b; Boyle et al., 1999; ECDG, 2012, World Bank Institute, 2011; McDonald et al., 2003).

<sup>10</sup> This was contrary to the principles of harmonization enumerated in several international treaties (e.g., the 2005 Paris Declaration on Aid Effectiveness, the 2008 Accra Agenda for Action, and the 2011 Busan Partnership Agreement) (Bamberger, 2009; Heider, 2011).

<sup>11</sup> "Evaluation must have a clear set of concepts to which all within an evaluation association can agree in principle" (Barbaric, 1999, p. 23). Therefore, if the lack of a common understanding on both the formal definition and the pragmatic articulation of ECB and ECD is adequately addressed, ECD is likely to enhance the effectiveness of ECD national evaluation associations.

<sup>12</sup> Grey literature included draft notes, conference presentations, short reports, and working papers that have been produced on ECD and ECB by a large number of international organizations (e.g., UNDP, UNICEF or the World Bank), research centers (e.g., IDRC, ODI, IDS), and in-country training institutions (e.g., the African Capacity Building Foundation).

noted by Preskill and Boyle in the conclusion of her article on the Multidisciplinary ECB Model recently published in the *American Journal of Evaluation* (AJE):

It is time to begin building a more robust knowledge base about ECB through empirical research. Our hope is that this model provides a jumping off point for designing studies on how these variables interact and affect the quality and sustainability of ECB efforts. There are many questions still to be answered; ECB is an area ripe for exploration (Preskill & Boyle, 2008, p. 457)

In order for this study to address such need, the first research question explored was:

- Question 1. To what extent is Evaluation Capacity Development (ECD) distinct from Evaluation Capacity Building (ECB) in international development contexts?<sup>13</sup>

Sub-question 1.1: What are the main central attributes of ECD as compared to those of ECB?

Sub-question 1.2: What, if any, relationship exists between the terms ECB and ECD?

The second factor affecting the effectiveness of ECD programming is the biased targeting of programs aimed at strengthening in-country evaluation capacity. On the one hand, national governments have received a privileged status within the scope of activities and programs aimed at enhancing national capacity funded by a number of international development partners since<sup>14</sup> the 1990s (referred to as either top-down, whole-of-government or enclave approaches)<sup>15</sup> (Lee, 1999; IOCE, 2006; Kusek & Rist, 2004; Rist &

---

<sup>13</sup> It is noteworthy mentioning that the difference between Capacity Building and Capacity Development does not really exist in the French language. The most frequently used expression in French literature is *Renforcement des capacités* and neither *construction* (equivalent to the term *building* in English) or *développement* (equivalent to the term *development* in English) is commonly used. For the purpose of this study, the term *renforcement des capacités* is to be considered equivalent to ECB. While the distinction between ECB and ECD might not be relevant to French-speaking contexts, the ECD model proposed by this study is still applicable to and potentially beneficial to ECD programming in all countries, regardless of their language.

<sup>14</sup> This was not surprising as the mandate of most UN agencies, for instance, is to support government and not CSO.

<sup>15</sup> Such emphasis on government-oriented capacity development originates, among others, from the 2002 Economic and Social Council resolution (ECOSOC resolution 2003/03) calling upon all UN agencies to

Stame, 2006; Unido, 1990; Wiesner, 1997). On the other hand, the number of non-governmental entities (civil society organizations, academia, and private sector)<sup>16</sup> involved in ECD programs has been particularly low (Leach et al., 2007; Mehta et al., 2010; Piper & Nadvi, 2010; Tembo, 2008).<sup>17</sup>

The result of such biased targeting has been two-fold. First, the technical skills and practices of evaluation professionals in countries have not improved<sup>18</sup> as much as they could have and the supply of local evaluation services has been discouraged, as attested to by not only the relatively low (although slowly growing) number of local firms providing evaluation services to government and Civil Society Organizations (CSO) but also by the less than optimal quality of evaluation deliverables produced.<sup>19</sup>

---

enhance development systems at country level by focusing on: (a) sustaining governments' efforts in developing country level strategies for capacity-building in the pursuit of internationally agreed development goals; (b) intensifying inter-agency information sharing on good practices and experience gained, results achieved, benchmarks and indicators, monitoring and evaluation criteria concerning capacity-building, and reflect them in the common country assessment and the United Nations Development Assistance Framework; and (c) inviting all organizations to include reporting on capacity-building in their annual reports to their respective governing bodies (UN, 2003).

<sup>16</sup> In a recent study using DAC bilateral data it was even found that almost half of the predicted value of aid was determined by donor-specific factors, one-third by needs, a sixth by self-interest and only 2% by performance (Hoeffler & Outram, 2008).

<sup>17</sup> The relatively marginal role assigned to what a growing number of practitioners nowadays refer to as VOPE (Voluntary Organizations of Professionals in Evaluation) is confirmed by the scant literature available on national evaluation associations' potential roles and responsibilities in the implementation of future ECD programs as well as the paucity of peer-reviewed studies on national evaluation associations' capacity needs. The dearth of specialized literature on these very specific ECD topics has certainly not provided an adequate platform for ECD national stakeholders to voice their concerns and needs in the area of ECD (Rocha Menocal & Sharma, 2008), and has discouraged further any possible systematic research effort to reframe ECD programming (including its assumptions and modalities of implementation). Only recently, some initiatives (still at an incipient stage) are being launched to alter the unbalanced ECD equation in favor of VOPE. Such is the case of EvalPartners, an innovative partnership established among the International Organization for Cooperation in Evaluation (IOCE), UNICEF and a variety of organizations, having as its primary objective to enhance CSO capacities to influence policy makers, public opinion and other key stakeholders, so as to inform public policies with evidence and ensure that equity and effectiveness play a central role in the development of country-led evaluation systems.

<sup>18</sup> The African Capacity Building Foundation (ACBF) was established in 1991 exactly to address the severe capacity needs identified among public institutions in Africa. Twenty-five years after its creation, ACBF (mainly sponsored by the African Development Bank, World Bank and UNDP) is still at the forefront of pan-African ECB and ECD initiatives. Thanks to its recent partnership with PACT, ACBF is increasingly working on the empowerment of CSO and the private sector towards better governance.

<sup>19</sup> Some meta-evaluations (UNICEF, 2004; CIDA 2007) conducted in the past, for instance, have shown that a third of evaluations are not worth their investment and another third are of uneven quality.

Second, activities aimed at strengthening technical capacity within national ministries and central planning agencies (Compton et al., 2001, 2002) have enhanced the knowledge of evaluation within the governmental sphere, but have not necessarily contributed to the development of skills to either formulate key evaluation questions or use evaluation findings (Bamberger, 2009; OECD, 2006a).<sup>20</sup> Furthermore, activities aimed at enhancing national evaluation capacity have rarely been customized to the specific functions (operational or strategic) and roles (commissioners, implementers, policy-makers) of individual officers operating within the government, and has instead favored the implementation of the same standardized approach at several levels within the government as if this were a monolithic bloc.<sup>21</sup> Addressing the limitation of the current ECS targeting is all the more relevant as the currently biased allocation of funding between governmental and non-governmental actors has three primary consequences. First, it has hindered the mainstreaming of evaluation at a more systemic level, as predicated by a number of studies, including a recent work funded by DfID (Gaventa & Barrett, 2010):

Change happens through multiple types of citizen engagement: not only through formal governance processes, even participatory ones, but also through associations and social movements that are not created by the state. Strengthening these broader social change processes, and their interactions, can in turn create opportunities for state reformers to respond to demands, build

---

<sup>20</sup> The limitations of capacity development approaches targeting national governments apply to areas other than evaluation, as well. The recent Global Monitoring Report (a joint World Bank/IMF reviewing countries' progress towards the attainment of MDGs), for instance concluded that the public sector capacity (including evaluation capacity) in the majority of developing countries has fared poorly against all MDG benchmarks (World Bank, 2012).

<sup>21</sup> The lack of specificity in the targeting of national governments is also due to limitations of two frameworks dominating the ECS discourse over the last two decades. The first one, represented by an obsolete supply-demand equation, stated that (i) the international development agencies were responsible for the supply to countries funds, trainings, and instructions on how to conduct evaluations; and (ii) the demand for evaluation services originated within recipient countries, often associated with the simple generic term of national governments or national institutions. The second framework was the one acknowledging the relevance of individuals, organizations, and the enabling environment in ECD programming, without providing an exhaustive definition of the different roles and functions to target within three levels, especially at the micro- and meso-levels.



external alliances and contribute to state responsiveness. (Gaventa & Barrett, 2010, p. 58)

Second, it has overlooked the fact that the identification of individual evaluation champions within host governments, characterized by high employee turnover, does not contribute, in most cases, to either the uptake of an evaluation culture or the sustainable promotion of evaluation use of findings in other sectors (Lennie, 2005). Third, it has ignored the development of an increasingly strong body of knowledge and skills among national evaluators, thus limiting the technical quality of their work and, therefore, their active involvement (e.g., as local evaluation team members) in donor-funded evaluations, an attestation of what some authors critically refer to as “elite domination” (Fung, 2003, p. 340). That notwithstanding, the targeting of ECD is already gradually evolving. The general interest in Voluntary Organizations of Professionals in Evaluation or Voluntary Organizations Promoting Evaluation (VOPE)<sup>22</sup> has increased over the last five years. That is more the case for development partners than national governments, as attested by 33% and 9%, respectively, of the 67 VOPEs surveyed in the course of a recent exploratory study (Holvoet et al., 2011). Therefore, a better understanding of how to make ECD targeting more inclusive might be beneficial to all the ECD practitioners, especially those who are currently confronted with the need for participatory strategies that have proved to work well in past ECD programs.

In order for this this critical issue to be addressed, the second research question addressed by this dissertation was:

- Question 2: To what extent could ECD targeting in international development contexts, become more inclusive in the future?

---

<sup>22</sup> VOPE are a fairly new phenomenon. As of September 2012, a total of 153 VOPE were identified: 127 national in scope and 26 with a more regional and international character (EvalPartners, 2012).

Sub-question 2.1: What is the current capacity of VOPEs?

Sub-question 2.2: To what extent are VOPEs currently involved in the conduct of evaluation and the promotion of an evaluative culture in their respective countries?

Sub-question 2.2: What are the factors characterizing the success or the failure of VOPEs that should be taken into account in view of their involvement in ECD programs in international development contexts in the future?

The third factor, hindering the effectiveness of ECD programming, is the lack of adequate operational tools specifically aimed at assessing the effectiveness of ECD as well as of VOPE activities and processes (Botcheva et al., 2002; Labin et al., 2012; OECD, 2006; Taylor and Clarke, 2008; World Bank, 2006, 2008). As a result, the implementation of ECD programs has often been based on unfounded and premature solution strategies (Lewis et al., 2006) rather than rigorous ECD needs assessments or either mid-term reviews or formative evaluations.

On the one hand, a recent global OECD assessment (OECD, 2010) attests to the lack of adequate capacity needs assessment (including in the area of evaluation) within the public sector in a number of countries:

What is missing from the relatively rich information about donor evaluation policies and the emerging evaluation capacities in the public sector in recipient countries is a clearer understanding of the links between these two. Information about donor use of country systems to evaluate development programmes is not readily available. Through the DAC Evaluation Network we are aware of the increasing involvement of governments in joint evaluation work, but there has been no stock taking of which parts of government are being involved and which country capacities exist to operate these collaborations on the partner country side.

On the other hand, the relatively generic indicators included in the Paris Declaration (OECD, 2005) and aimed at measuring the effectiveness of capacity development programs,

including those implemented in the area of evaluation, confirm the limitations of the evaluation tools currently in use to assess ECD effectiveness (Table 1).

Table 1

*Limitations in Measuring CD Internationally: The Case of the Paris Declaration on Aid Effectiveness*

Indicator	Definition
Strengthen capacity by coordinated support	Percent of donor support provided through coordinated programmes consistent with partners' national development strategies
Strengthen capacity by avoiding parallel structures	Number of parallel project implementation units per country

Source: OECD, 2005.

There is an obvious need for tools and indicators that could better capture both the current in-country evaluation capacity needs and progress that is likely to be attained as a result of the participation in ECD programming. Identification or development of effective ECD metrics is all the more relevant if one takes into account not only the findings of a recent assessment that reports that only three of the 54 countries surveyed had results-oriented frameworks that were deemed adequate (OECD, 2008), but also the conclusions of a more recent work, according to which:

Capacity building-strategies and plans, including a diagnosis of the situation and a needs assessment, are generally not developed as a component of the M&E implementation. The result is little knowledge of potential weaknesses and areas of improvements in M&E capacity and an inadequate design of capacity-building activities, which often look like a long list of uncoordinated short-term capacity building activities (Clotteau et al., 2011, p.175).

In order for this this third issue to be addressed, the third research question addressed by this dissertation was:

- Question 3: How can, or how should, ECD be best evaluated?

Sub-question 3.1. What are the key criteria that need to be taken into account in order to assess ECD effectiveness?

Sub-question 3.2. At what level do identified key ECD criteria need to be evaluated?

### Aim and Scope of the Dissertation

By acknowledging the existing knowledge gaps in both international ECD practice and research, and with the objective of providing donors with some actionable recommendations on how to enhance their respective ECD programs in the future, this dissertation seeks to address the three key research questions presented above.

First, in order to fill the “definitional gap” (what makes capacity “development” different from capacity “building”), it was intended through this study that a new and clearer definition of ECD would be developed, with a special focus on its central attributes, especially if compared to ECB and other activities aimed at strengthening national evaluation capacities in international development contexts. A common understanding of ECD was believed to add value to the current ECD discourse as it would not only contribute to enhancing donors’ harmonization in ECD programming but it would also allow a more rational allocation of corresponding ECD resources in the future. The clarification of the relationship between ECD and ECB was also believed to restore the essence of ECD, often turned into a neutral, value-free catch-all term incorporating any form of technical assistance and training made available in international development contexts. In doing so, great effort was made in attempting not to turn this endeavor into an academic exercise focused on semantic issues. Mindful of the needs and interest of hundreds of evaluation professionals

working in sub-Saharan Africa, Latin America, and South-East Asia, this study was based on the views and opinions expressed by practitioners on this issue. Furthermore, one additional objective pursued by this study was to voice the needs and interests expressed by many key ECD stakeholders who are not always in a position to sit at the table with the ECD commissioners and funders.

Second, in order to address the limitation of ECD targeting, and especially the less than optimal involvement of representatives from civil society, the study was aimed at enabling a critical shift in ECD programming: from the long-established vision of national evaluation associations or VOPEs as being merely ECD program recipients to a new understanding that VOPEs could be “engageable” ECD service providers and active promoters of an in-country evaluative culture (Constantinou, 2007). In order to do so, a framework (referred to as the SFAR Framework presented in Chapter IV) was developed that shows the key ECD stakeholders that ought to be targeted at the national level, as well as the interactions existing among them that ought to be supported and facilitated. The framework was tested and enhanced based on a review of the ECD ecology in three different countries (Democratic Republic of Congo, Niger and South Africa). Exploring the extent to which VOPEs could enhance ECD appeared all the more timely, as a large number of them were currently developing their missions and programs in view of their professionalization at the time the study was being conducted. As a result, the emphasis of this study was not to look at VOPEs in isolation but rather in terms of their purpose and relationship within the world where they functioned and interacted with other evaluation stakeholders, as predicated by a variety of evaluation scholars who have embraced the systemic thinking paradigm (Kurtz & Snowden, 2003; Morell, 2010; Patton, 2011; Ramage & Shipp, 2009; Snowden & Boone, 2007; Williams & Hummelbrunner, 2010).

In order to address the lack of adequate tools to evaluate the effectiveness of ECD endeavors, this study aimed at providing a few tools that would assist donors and national governments in the evaluation of ECD initiatives that they would be able to use in the future. Such tools were intended to enhance the design and implementation of more “democratic” and “participatory” ECD programs in international contexts, with a special emphasis on ownership and sustainability. This seemed particularly timely as ECD is gradually assuming the characteristics of a distinct area of evaluation practice.

Overall, building on the feedback provided by a number of purposefully selected ECD scholars and practitioners interviewed globally, this study was expected to provide donors with a fairly comprehensive set of definitional, conceptual, and operational tools that they might want to use to inform their ECD programming in the future. In its ambitious and yet humble intent, this dissertation calls upon all ECD practitioners to take into account the two key principles of ownership and donor harmonization, mentioned in the Paris Declaration of Aid Effectiveness (OECD, 2005), the Agenda for Action (2008) and the country-led “Delivering as One” initiative at the UN (Heider, 2011).

In an effort to align itself with the dialogue currently taking place within the international evaluation community, validation of preliminary findings and conclusions was sought among a variety of donors, leading evaluators, and VOPEs opinion leaders, including: (i) Directors of Evaluation offices and evaluation officers in a number of international organizations, such as, the World Bank, the African Development Bank, UNICEF, and the United Nations Development Programs; (ii) Evaluation officers at Foundations, such as, the Rockefeller Foundation; (iii) ECD champions within national governments; (v) active members on the International Development Evaluation Association (IDEAS) Listserv; (vii) qualified international development evaluation practitioners working

for consulting firms in the Washington DC area; and (viii) representatives of the IOCE Board of Directors and members of other VOPEs.

By addressing the three key research questions presented above, this dissertation was sought to inform the design of, and the resource allocation for, future donor-funded ECD activities and programs in a variety of countries. Furthermore, in fostering the participation of a plurality of ECD stakeholders with different agendas and perspectives, this dissertation had the ambition to both serve a catalytic function and provide a comprehensive list of ECD-related topics that researchers might want to explore in the future.

### Organization of the Dissertation

This dissertation consists of five chapters. In the current chapter, the three main gaps in the current ECD discourse as well as the corresponding research questions that this study was aimed at addressing were presented.

In Chapter II, the findings of a literature review focused on three main themes (the main theoretical frameworks dominating the ECD discourse to this date, the biased targeting of ECD programming, the most common variables used to evaluate ECD) are presented. Through the analysis of the strengths and weaknesses associated with each of the theoretical and evaluative frameworks presented in this chapter, the foundation is laid for the development a more robust and comprehensive framework for both conducting and evaluating ECD that will be tested in the field.

In Chapter III, the overall design, as well as the specific sampling strategy and data collection tools used, to address the three key research questions are presented.

In Chapter IV, the results of both the systematic literature review and field data collection are presented. In the first section, a list of definitions of ECD-related terms,

presented in chronological order and commented on one-by-one, is provided. In the second section, three case studies on VOPEs from three different countries (sub-Saharan Africa) are presented. In the third section, the most common variables used in the field to assess ECD effectiveness, as well as some examples of VOPE theories of change are described.

In Chapter V, the conclusions of the study are presented in the first section and the answers to each of the three key research questions and sub-questions are provided. The limitations of the study as well as its implication for future ECD research, theory, and practice are also discussed in the second section of the chapter.



## CHAPTER 2

### REVIEW OF THE LITERATURE

Borrowed from the title of a popular evaluation book<sup>1</sup> (Rist & Stame, 2006), the idea of “shifting from studies to streams”<sup>23</sup> represented one of the foundational blocks of this study for two main reasons. First, the belief, adequately supported by empirical data, that individual practitioners doing evaluation (more than written articles or reports that discuss evaluation) provide a critical contribution to the current state of professional evaluation practices at the global level. Second, and in line with the tenets of the collaborative immersion approach<sup>24</sup> (Hufmann et al., 2008), the assumed need for assessing the link between individual capacity and organizational growth. However, wary of the multiple dynamics associated with systems where evaluation functions are embedded, this study was also based on the assumption that individual evaluation capacity is necessary for organizations to grow, but is, in itself, not sufficient to ensure growth (Douglass et al., 2003).

---

<sup>23</sup> According to Rist & Stame, real time streams of evaluation knowledge are needed, as opposed to individual (formative and summative) studies

<sup>24</sup> Based on real-world evaluation experiences, this approach is similar to empowerment and participatory evaluation but, differently from these two, it has as primary objective, to enhance learning and not simply to conduct an evaluation.

## Introduction

A review of the most relevant ECD, ECB and VOPE literature currently available is presented in this chapter. The review of literature includes both peer-reviewed journal articles and work conducted and disseminated by seasoned practitioners on a number of ECB- and ECD-related topics in Asia, Latin America and sub-Saharan Africa.

The utility of this literature review was double-fold. First, by highlighting the theoretical contributions made by several international development scholars to the definition, interpretation, implementation and evaluation of ECD and VOPE over the last two decades, the review situates the study within the broader international development evaluation discourse. Second, by both examining the innovative features and pitfalls of past and current ECB and ECD definitions (including their respective targeting strategies), and recognizing the distinctive ideological properties of ECD and ECB frameworks developed in the past (which the implementation and evaluation of a plethora of corresponding initiatives had been influenced by for years), this review was aimed at facilitating a renewed understanding of both ECB and ECD. Based on the premise that this study was not to be a stand-alone work on ECD, this literature review included the work of numerous evaluation practitioners and theorists who, over the years, had engaged in empirical studies as well as epistemological debates over not only ECD, but also ECB and CD more in general<sup>25</sup>.

This chapter consists of three main sections. In the first section, more definitional in nature, the current dilemma over the use of ECB and ECD in both theory and practice is

---

<sup>25</sup> This is one of the key assumptions, which this study rests on. Given the paucity of peer-reviewed literature on ECD, some of the constructs developed in the CD field will be adapted and used as needed to characterize ECD processes within the scope of this study.

presented<sup>26</sup>. In the second section, an overview of the frameworks used over the last 15 years to describe and interpret both ECB and ECD processes and targets (including the corresponding stakeholders and influencing factors) is provided. In the third section, which was primarily aimed at informing the measurement of ECD effectiveness in the future, a number of frameworks used by international development agencies, research centers, and national governments to evaluate ECB and ECD processes and effectiveness, are discussed.

Overall, the primary objective of this chapter remains to describe the definitional and conceptual foundations of ECD that the study will build on, with the objective of moving the current discourse on ECD forward. In particular, the objective of this chapter is to provide in-country ECD stakeholders and development partners with both a better ECD epistemological framework and a set of practical tools to enhance the effectiveness of ECD programming in the future.

#### The Current State of ECB and ECD Literature

The body of the existing peer-reviewed literature on ECB and ECD, both within and outside the area of international development, is somewhat limited. The very limited number of results yielded by a cursory search for both terms “Evaluation Capacity Development” and “Evaluation Capacity Building” on the Social Science Citation Index on August 1, 2012 is a clear illustration of the limited peer-reviewed literature on the subject<sup>27</sup>.

By contrast, a copious quantity of draft notes, conference presentations, short reports and working papers on ECB and ECD has been produced by a large number of

---

<sup>26</sup> This overview is necessary and useful to better understand the typical attributes of ECD and how it relates to ECB and other types of activities aimed at enhancing in-country evaluation capacity. In doing so, the review of ECD key features built on the commonalities of the numerous definitions currently available and reconcile some of the existing differences.

<sup>27</sup> The search for ECD did not yield any result and the one for ECB yielded 31 results (not all of them relevant to this study).

international organizations (e.g., the United Nations Development Program, the United Nations Children’s Fund, the World Bank), research centers (e.g., the International Development Research Centre, the Overseas Development Institute, the Institute of Development Studies at the University of Essex) and in-country or regional training institutions (e.g., the African Capacity Building Foundation) over the last decade.

The popularity of both ECB and ECD within the international evaluation community was confirmed by the results of a content analysis (conducted concurrently with the first phase of the literature review) of online discussions taking place among members of several online Evaluation Communities of Practice (ECoP) between 2011 and 2012.<sup>28</sup> Based on this analysis, ECD appeared to be more popular in international development evaluation list-serves (such as XCeval)<sup>29</sup> than EVALTALK, the more general list-serve sponsored by the American Evaluation Association (AEA). Respectively, 187 and 36 postings were retrieved through a search for “Evaluation Capacity Development” in the two ECoP archives. In contrast, ECB seemed to be a more recurrent topic in non-international development-focused list-serves. For instance 50 results were yielded by a cursory search for “Evaluation Capacity Building” in the EVALTALK archive, which corresponded to 24 more results than those generated through an identical search for exactly the same term in the XCeval archive.

At first the speculation over why there was such a difference in the use of the two terms (that is, the apparently wider use of ECD in international development contexts and the more frequent adoption of the terms ECB in sectors outside international development),

---

<sup>28</sup> The content analysis was conducted by the study’s author before the study questions were developed.

<sup>29</sup> The XCeval list-serve, originally affiliated with the International and Cross-Cultural Evaluation Topical Interest Group (TIG) at the American Evaluation Association (AEA), is primarily aimed at practitioners with a keen interest in international development evaluation.

seemed a plausible one. However, later on, the apparent lack of a clear rationale for systematically preferring one term (ECB) over the other (ECD) among individual international development practitioners<sup>30</sup>, as emerged by a series of semi-structured interviews with ECD scholars, seemed to suggest otherwise<sup>31</sup>.

That a certain confusion or lack of consensus existed over the meaning of both terms was attested by a certain stream of peer-reviewed literature produced by evaluation scholars reflecting upon their own practice and that of their colleagues. Among the most recent contributions on this topic, Nielsen & Attström's (2011) appeared particularly relevant. According to both Danish authors, a more serious reflection and debate on the distinction between ECD and ECB was particularly needed as it would allow addressing four main issues affecting evaluators' practice in a number of countries. First, the widespread conceptual pluralism in the area of ECB and ECD. Second, the increased number of discordant opinions regarding ECB purposes. Third, the lack of a comprehensive empirical base for most ECB and ECD models. Fourth, the relatively greater focus on approaches implemented to tackle CB rather than CD.

Supported by the scholarly work consulted during the first phase of the literature review, one of the central assumptions underlying this study—that ECB and ECD were not completely identical terms—was generally supported. Far from remaining an isolated

---

<sup>30</sup> The tendency among many practitioners and researchers to use ECD and ECB interchangeably seems to discourage any effort to recognize any distinctive and unique meaning to either term. The recent synthesis on the ECB empirical literature recently published on AJE (Labin et al., 2012) certainly represents a commendable effort in systemizing and codifying the key elements of a term whose meaning has evolved over time and has been characterized by different connotations depending on the context where it was used. However, the association of ECB and ECD made by the authors as if they meant exactly the same thing, without making this assumption explicit, does not seem to hold in the case of international development contexts, where ECD (rather than ECB) is increasingly considered a critical factor in the promotion of an evaluative thinking at the institutional and not merely individual or organizational levels.

<sup>31</sup> That said, a certain number of agencies, including OECD and the Independent Evaluation Group at the World Bank have repeatedly used the terms ECD rather than ECB over the last few years.

impression, the existence of different paradigms associated with ECB and ECD also was confirmed by the recognition among governmental agencies of both CB and CD as two different, although complementary, areas of work (e.g., the Office of Capacity Building and Development within the U.S. Department of Agriculture)<sup>32</sup>.

Based on the growing interest in the clarification of ECB and ECD central attributes highlighted by this literature review and in response to the four main issues mentioned by Nielsen and Attström (2011), this study sought to provide an informed contribution to the existing debate on the relationship between ECB and ECD. That appeared all the more relevant as the in-depth exploration and identification of the commonalities and differences between the two terms, as perceived by in-country practitioners involved in this study, would make it possible to recognize the limitations of the peer-reviewed research currently available on this topic and would help filling the gap between theory and practice. In addition, any further reflection on the ECB/ECD dilemma introduced in this section and developed further in the next two chapters (e.g., Are the two terms the same or different? And, if they are different, what are their distinctive patterns?) was believed to give voice to those practitioners in international development contexts who endorse a distinction between ECB and ECD on the grounds of different political and ideological premises (CDRA, 2007).

### The ECB and ECD Definitional Dilemma

ECB and ECD have become part of evaluation practitioners' everyday language since the late 1990s. However, despite the increasing use of both terms over the last decade,

---

<sup>32</sup>The Office of Capacity Building and Development (OCBD) is currently involved in a variety of international projects (USDA, 2011). In 2012, for instance, the OCBD Monitoring and Evaluation Staff (M&ES) started working very closely with the Fragile Market Economies Division (FMED), a newly established division within USDA, and provides it with a number of services in both implementing and integrating the Results Oriented Management (ROM) Monitoring and Evaluation System into their current management practices.

a precise and agreed upon definition that could describe unambiguously their respective attributes was still lacking in the late 2000s, as stated in a few articles on ECB:

The very concept of ECB may be liberally operationalized in various organizational milieus, especially since there is still no consensus among researchers or practitioners on an operational definition of ECB (Volkov, 2008, p. 193).

More effort should be made to further explore and develop the concept of ECB. The need to assess the value of different definitions of ECB becomes of greater concern for us to be able to build consensus “around its socially constructed meaning in an iterative fashion” (Taut, 2007, p. 120) (Volkov, 2008, p. 195).

The apparent paucity of definitional and conceptual efforts aimed at filling this void (e.g., by clarifying the commonalities and differences between the two terms) came as a surprise as several authors had lamented the inherent complexity and vagueness of the ECB and ECD constructs (and therefore implied the need for addressing such gap) already in the late 1990s. When referring to CB, for instance, Morgan (one of the most prolific authors on capacity), defined it as:

...A risky, murky, messy business, with unpredictable and unquantifiable outcomes, uncertain methodologies, contested objectives, many unintended consequences, little credit to its champions and long time lags (Morgan, 1998, p.6).

Likewise, in defining CD, Lusthaus (one of the most well-respected Canadian experts in institutional evaluation and change) defined it as:

...A concept still in its infancy. Its definition is still forming. Research describing how people use the concept is sparse. So is research, which tests its assumptions and predicts its consequences. There are few evaluations of projects that are claiming to use approaches to capacity development (Lusthaus et al., 2002, p.34).

That notwithstanding, the popularity of both terms had been unquestioned for over two decades. The fact that the themes of the 2001 and 2002 AEA annual conferences were “Evaluation Capacity Building” and “Mainstreaming Evaluation” respectively, certainly demonstrated the relevance and public recognition of both terms. Likewise, the fact that 73 sessions and posters on ECB and ECD (from a total of 1,000) were delivered during the 2012 AEA conference is another attestation of the enduring popularity of the two terms.

However, between the two terms, peer-reviewed literature focuses more extensively on ECB for a variety of reasons. First, the integration of ECB in a number of widely adopted participatory evaluation approaches (Greene, 2005),<sup>33</sup> has attested to both the acceptance and formal recognition of ECB ontological premises and values within the evaluation community, especially in the U.S., but fails to give sufficient credit to ECD<sup>34</sup>. Second, the findings of a 2008 survey administered among AEA members (Preskill & Boyle, 2008), according to which half of respondents were engaged in ECB efforts within the scope of their respective work, had reconfirmed the centrality of ECB but not of ECD, despite the fact that the evaluation offices of several agencies and multilateral organizations (e.g. the World Bank and OECD) had been using ECD over ECB for almost a decade.

---

<sup>33</sup> These include the following: a) collaborative evaluation (O’ Sullivan, 2004; Rodriguez-Campos, 2005); b) empowerment evaluation (Fetterman & Wandersman, 2005); c) internal evaluation (Love, 2006); and d) participatory evaluation (Cousins & Whitmore, 1998).

<sup>34</sup> Despite the lack of specific work on values in international development evaluation guiding principles in evaluation are gaining increasing recognition. As Heider (Heider, 2011) illustrates in a very effective manner in her work on the evaluation of ECD<sup>34</sup>, the framework for peer reviews of evaluation functions in the UN developed in 2007 identified three core principles of the evaluation profession: independence, credibility and utility (DAC/UNEG 2007). The first principle, independence, often associated with impartiality, is well aligned with the three-level ECD framework, in that it could be attained at the individual (what is normally referred to as training individuals), organizational (institutional development) and institutional level (capacity development). The second principle, credibility, is enhanced by impartiality but it also requires competent evaluators and transparent evaluation processes. The third principle, utility, is enhanced by the intentional (Baizerman et al., 2005) and timely use of evaluation findings for decision- or policy-making and assumes the accessibility of evaluations.



As a result, there appears to be a paradox confronting scholars and practitioners, that is, the limited inclusion of ECD in peer-reviewed research produced in both Canada and the U.S., despite the frequent use of the term both among evaluation practitioners and development partners in international development contexts. Such a paradox deserved being explored further as it was exacerbated by an ideological dilemma emerging within the evaluation community: the interchangeable use of ECB and ECD among a larger number of US-based evaluation practitioners (King, 2002) as well as the inevitable practical programmatic implications of such fluid use of the two terms.

In particular, ECB seemed to suggest that donor-funded interventions were both one-time and unilateral efforts aimed at starting evaluation capacity from scratch within countries based on a pre-conceived design (zero-base approach). Such definition echoed two main arguments that other scholars had put forward in the past (Chambers, 1997; CDRA, 2007; McAllister, 2011; Nelson et al., 2009). First, that ECB did not always do justice to the knowledge and wealth of evaluation experience already existing in countries targeted by ECB programs (though not always systematically organized) (Fukuda-Parr et al., 2002)<sup>35</sup>. Second, that that assumption according to which short-term injections of know-how (normally associated with the concept of ECB) would address the weaknesses identified within evaluation systems, no longer held<sup>36</sup>. Surely, the often-limited effectiveness of short-term trainings, one of the most recurrent critiques to ECB, was not an unprecedented realization.

As Toulemonde put it:

---

<sup>35</sup> This idea resonates with Freire's argument that faith in people is precondition for dialogue (Freire, 1970).

<sup>36</sup> Such short-term efforts were even interpreted at times as an impediment to CSO development. Co-opted into donors and governments' agenda (Schaumberg-Muller, 1996) and provided with technical assistance to fare better in areas imposed from the outside, CSO fell prey to the "NGO-ism" phenomenon and got involved in the "empowerment by privatization" model (Yachkashi, 2010).

These events rarely last for more than one or two days. They can hardly be considered as actual training and should be qualified as initiation. They give grounding in evaluation, increasing knowledge, but are not a substantive investment in skills development (Toulemonde, 1995 p. 23).

Similarly, the United Nations (UN) already in the early 1990s had acknowledged that the ECB focus on individual training was not effective as initially imagined, as individuals did not act independently or isolated from their context but operated within the context of larger organizations. Therefore, in response to such realization, trainings started being delivered to a larger number of staff within the same organization. In addition, unprecedented discussions on organizational issues were initiated in the course of these very same learning events (UNIDO, 1990).

Furthermore, ECD appeared to be distinct from ECB given its unique political and ideological connotation. That is, its stronger emphasis on inclusiveness, flexibility, development result-focus and context-responsiveness to an already existing evaluation capacity. If applied to the evaluation context, the distinction between CB and CD, discussed in a recent report (Ortiz & Taylor, 2009) commissioned by the United National Educational, Scientific and Cultural Organization (UNESCO) to the Institute of Development Studies (IDS), contributed to confirming and amplifying the differences between the two terms. As discussed in one of the report's most salient passages, CB and CD (and therefore, ECB and ECD) were not simply described as two different terms but rather as two opposite development paradigms:

Much of the capacity development literature stresses the fact that development is already happening before the arrival of any project, donor, program 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:

- a) 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;
- b) They do not build on momentum; that is, positive development initiatives and processes already in motion;
- c) 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. 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 behavior changes are unlikely to develop deep roots (Ortiz & Taylor, 2009; p. 26).

As a result of the ECB and ECD definitional and epistemological gaps existing in peer-reviewed literature, grey literature and evaluation practice (Table 2), an informed consensus on how ECB and ECD relate to each other was still missing. Likewise, provided that differences existed between the two terms, it was not clear yet how differently ECB and ECD needed to be implemented and evaluated.

Table 2

*ECB and ECD Definition and Hermeneutical Gaps*

ECB STAKEHOLDER	LIMITATIONS
ECD SCHOLARS	Lack of articles on ECD in peer-reviewed literature
RESEARCHERS AND DEVELOPMENTS PARTNERS	Contrast between: <ul style="list-style-type: none"> <li>- The interchangeability of the two terms in some of the grey literature available and</li> <li>- The preference of one term over the other in some other grey literature</li> </ul>
IN-COUNTRY PRACTITIONERS	Contrast between: <ul style="list-style-type: none"> <li>- The identification of an ideological distinction between ECB and ECD among a certain number of in-country development evaluation practitioners and</li> <li>- The interchangeability of the two terms among some other in-country development evaluation practitioners</li> </ul>

## Definitions of Terminology

For the purposes of this study, the following terms were used (the new definitions of ECB and ECD, which were developed as a result of the field data collection and analysis, are presented in Chapter V).

Advisory Panel: A group of ECD specialists providing advice during both the development of the study proposal and the validation of the preliminary findings (three phase process). The Advisory Panel consisted of three sub-groups: (a) ECD leading scholars; (b) VOPE Coordinators and VOPE members in the three countries included in the case studies and referred to as “in-country practitioners”; and (c) ECD officers working within development agencies and ECD specialists working in countries other than those included in the three case studies

Evaluation Capacity Building (ECB): Mostly associated with short-term activities, ECB refers to building the skills and ability (human capital) of individual evaluators around the world using systematic research methods to evaluate the performance of projects, programs, country development strategies, and global programs<sup>37</sup>.

Evaluation Capacity Development (ECD):

The process whereby people, organizations and society as a whole unleash, strengthen, create, adapt and maintain evaluation capacity over time<sup>38</sup>

---

<sup>37</sup> This definition was adapted from the one developed by the former director of the World Bank (WB) Operational Evaluation Department (the predecessor of the current independent Evaluation Group) (McAllister, 2011, p.214)

<sup>38</sup> This definition was adapted from the OECD definition of capacity development (OED, 2006, p. 12).

Voluntary Organizations of Professionals in Evaluation (VOPE): VOPEs include formally constituted associations or societies, as well as informal networks. Their membership is open to not only those who conduct evaluations but also to those who commission and utilize evaluations and those engaged in building the evaluation field <sup>39</sup>.

## Inclusiveness of ECB and ECD Targeting in International Development Contexts

### Introduction

In order to understand the relationship existing among ECB and ECD stakeholders and as a way to assess the degree of equity and inclusiveness inherent to ECB and ECD processes (e.g., the extent to which VOPE are actively involved in the planning, implementation and evaluation of ECD efforts) (Box 1), the literature available on four key topics will be reviewed in this section. In the first part, an overview of the general inquiry framework underlying the study of both VOPE capacity and their involvement in ECD programming, will be introduced. In the second part, the evolution of ECD targeting<sup>40</sup> and the processes, inherent to the institutionalization (Stufflebeam, 2002) of the evaluation function in a variety of international development contexts over the last two decades, will be explained. In the third part, the main ECB and ECD frameworks dominating the international development discourse over the last two decades will be presented. In the

---

<sup>39</sup> This definition corresponded to the official one presented in the EvalPartners website (EvalPartners, 2012). Link: <http://www.mymande.org/evalpartners/the-international-evaluation-partnership-initiatives> (accessed on Nov. 2, 2012).

<sup>40</sup> Targeting is not only about selecting individuals who are likely to benefit from the intervention being planned (as mainly referred to in this chapter) but also about identifying areas where the intervention will take place. In the case of ECD, this is a particularly sensitive topic. As attested five case studies recently carried out in five countries (Cameroon, Montenegro, Peru, Tanzania, Vietnam) (Ubels, 2009), the supply of ECD services has been criticized for discriminating against rural areas (local demand is not met by CB programmes implemented at the national level and the fees imposed on sub-national actors for accessing CB programs are often prohibitive) as well as for using standardized (rather than tailored) CB approach and not benefiting from the contribution of sub-national actors (other providers) whose access to the ECB market is very limited.

fourth part, a review of the evolving role played by VOPE in the promotion of an evaluative culture in a number of development contexts will be provided.

### General Inquiry Framework

In addressing the second research question on the degree of inclusiveness associated with the targeting of current ECD programming, this study was grounded in a small number of theoretical frameworks stressing the critical role that individuals have not only in creating and sharing knowledge but also in shaping professional practices and influencing reality around them. Consistent with the constructivist approach (Piaget, 1995; Bruner, 1996; Guba & Lincoln, 1985), this study particularly emphasized the role played by VOPE in both the development and dissemination of evaluation knowledge and methods. Aligned with the growing body of literature asserting that human knowledge, when generated by—rather than imparted to—individuals is key to promoting learning and change (Bawden, 2008), this study was aimed at exploring feasible opportunities for more actively involving VOPE in ECB and ECD programming, not only as recipients but also as providers.

Based on the premise that more support for VOPE was needed for a more sustained proliferation of an evaluation culture in a variety of countries, this study recognized the lack of adequate entry points for national evaluators to inform the current debates on international development evaluation and, more specifically, on ECD. In so doing, this study was well embedded within the equity-based discourse in international development (Reynolds & Williams, 2012; Bamberger & Segone, 2011).

First, the study drew on the Change Collaboration Theory (Himmelman 1996, 2001), which focused on the critical role played by ownership and power within the scope of any collaborative relationship. In particular, one of Himmelman's ideas that had a tremendous

influence on the design of this study was that every collaboration (such as the one between donors and aid recipients) was not characterized by a status quo (Yachkashi, 2010) but rather by a fluid power sharing and the possibility for continued transformation. As a result, the relationship between development partners and host countries, which had often been described in terms of a supply and demand equation (funder having a more prominent role as funding suppliers over those who demand their intervention and “consume” aid resources), was presented in this study not as a static interaction, but a condition that could be altered.

Besides Hilleman’s Change Collaboration Theory, this study was also inspired by Hay’s work on participatory evaluation (Hay, 2010). In particular her thoughts on how to get local ECD stakeholders more involved in the current discourse on evaluation methods and practice and the critical analysis of her field experience in South Asia, represented a tremendous contribution to the conceptualization of this study:

“[...] One tremendous opportunity in South Asia to advance the evaluation field is the history of praxis and engagement in development research and programming in this region. At the implementation or grassroots level, and among resource groups supporting grass roots development, there are thousands of people and groups engaged in and learning from innovative and contextualized development work. The ground level experience of such practitioners can help situate evaluation theory, methods, and application within a framework of use and practice” (Hay, 2010, p.189).

### The Biased Targeting of ECD and ECB Programming

Most of the ECB literature published between 1990 and 2010 focused on the opportunities and challenges associated with the strengthening of governments’ evaluation capacity (Smits, 2011). This was in line with the popular economic argument that evaluation was an unintentional and environmentally induced system and that, therefore, institutions were the key players in the promotion of a national evaluative culture. Consistent with this

paradigm, a number of authors (Bates, 1995; Derlien, 1990; Wiesner 1993,1997) had argued that the value of evaluation in international development contexts lied in its ability to both unveil all the inefficiencies of the public sector and provide tax payers and civil society with timely information on successes and failures of public programs. Otherwise said, the role of the evaluations conducted within the public sector (e.g., contributing to greater government's effectiveness) was viewed as equivalent to that of prices and competition in commercial markets (e.g., determining the profitability of businesses in the private sector).

The link between ECD and disciplines other than evaluation was not a new one. Already in the past, inspired by institutional economics and the principles of information theory (Krippendorff, 1986) and incentives (Lahiri, 2007; Podems, 2007), several scholars had called upon development partners to implement ECD initiatives specifically aimed at national governments as a way to enhance the latter's welfare-enhancing power (Rist & Stame, 2006). Building on the work of these earlier authors, development partners funding ECD programs around the world quickly embraced the argument that it was necessary to first create "educated consumers" in evaluation among policy makers and users of evaluation (Morris, 1994)<sup>41</sup> and only after having received their buy in, target entities outside of the governmental sphere VOPE (see Box 1).

As a result of the development community's buy-in, a large amount of resources were allocated by donors in a variety of countries around the world since the 1990s, with the specific goal to foster both the demand for and the use of evaluation among national

---

<sup>41</sup> However, rather than confining policy-makers' and decision makers' understanding to evaluation accountability purposes only, Morris, stresses the relevance of making them familiar with the evaluation learning purposes, as well.



governments (Boyle & Lemaire, 1999)<sup>42</sup>, as part of what began being described as a renewed social contract (World Bank, 2006). Development partners' responses were not a naïve effort to fill evaluation capacity gaps observed in the countries that they supported, but rather reflected of a solid body of literature arguing in favor of governments' involvement in

### **Box 1. Basics about VOPE**

- Two of the oldest evaluation societies (the Malaysian Evaluation Society and the Ghana Evaluation Association) were founded in 1995 and 1997 respectively (Ghana Evaluators Association) (p.16).
- The number of evaluation seems to have risen since 2004, under the influence of regional evaluation societies were created (e.g., AfrEA established in 1999) and, more importantly, of IOCE in 2003 (Segone & Ocampo, 2006). An alternative explanation might be that in the realm of the Paris Declaration, some bilateral donors as well as multilateral organizations (such as UNICEF) started to give more attention to the establishment and use of country-led M&E systems (p. 9).
- Nearly 75% of VOPEs do not have any paid staff at their disposal. Those who do have less than five. Seventy percent of the evaluation societies benefit from the support of volunteers (the number ranges between 1 and 10 volunteers (p.17).
- Forty percent of the evaluation societies have no financial means at their disposal. Of the other sixty percent that do dispose of financial means, about half has less than 10,000 U.S. Dollars (p.20).
- Although most associations rely on membership fees as their main source of income, the second most important source of funding is provided by international donors (p. 20)
- Almost half of the evaluation societies organize a meeting more than twice a year (p.24).

Adapted from: Holvoet, N., Dewachter, S., Gildemyn, M. (2011)

---

<sup>42</sup> Investing in program enhancing government's evaluation capacity has been quite fruitful in several but not all cases. As the experience shows in a number of countries (e.g., China, Columbia, Indonesia, Mexico), it is only when governments become aware of the evaluation purposes and start feeling the public pressure to deliver better services thus developing a demand for evaluation, that the institutionalization of the evaluation function becomes effective. A well-articulated ECD programming is certainly key to mainstreaming the evaluation function in the government's agenda (Sanders, 2003). However, as the case of Columbia and Mexico demonstrate, for ECD to be effective, a context-responsive planning will be particularly needed.

national ECD programming. Through a review of the specialized literature from the late 1990s, focused on the role played by national jurisdictions attitudes towards evaluation over the years were traced (Derlien, 1990; Gray, Jenkins & Segsworth, 1993; Leeuw, Rist & Sonnichsen 1994), and three main waves were identified (Toulemonde and Rieper 1997; Mayne and Zapico-Goñi 1997; Bemelmans-Videc, Rist & Vedung 1997). During the first wave (1960s), evaluation was strongly centralized and primarily aimed at improving the performance of public programs (also known as the “Great wave”) in such countries as the U.S., Canada, Sweden, and Germany. During the Second Wave (1970s and 1980s), the evaluation function was mostly deferred to parliaments and used for accountability purposes in such countries as Denmark, Norway, the Netherland, Great Britain, Finland and France. During the Third Wave (1990s), as the evaluation function was shifting back to governments’ executive branches (Colombia, Ireland, Korea, Indonesia, Switzerland, all countries in the European Union as a result of the Structural Funds requirements), an increased interest in institutionalizing it quickly developed.

The rationale for more ECD programs specifically targeting national governments was that the more evaluation was going to mainstreamed within the governance structure<sup>43</sup>, the more successful its implementation would be (Boyle & Lemaire, 1999; EuropeAid, 2008; UNEG, 2012). However, the practice of assigning such a prominent role to national governments in ECD programming (the so-called demand side of the ECD equation) had some unexpected consequences. It inevitably discouraged the implementation of activities and programs enhancing the evaluation capacity of non-institutional entities (e.g., civil society organizations and local evaluation associations). When acknowledging that

---

<sup>43</sup> Three countries (and their respective ECD targets) where the efforts made to institutionalize evaluation were particularly successful include Colombia (Controller General), China (Auditor General) and Indonesia (National Development Planning Agency) (Guerrero, 1999).

“evaluative knowledge was the public sector-manager’s surrogate for what was known as profit and loss statement in the business sector” (Havens, 1992, p. 34), the biased targeting of ECD programming ended up ignoring the relevant oversight role that civil society could have played<sup>44</sup>. To use Haven’s metaphor, public managers (e.g., government officials) receiving information and training on evaluation failed to sit at the table and discuss with the Board of Directors (e.g., citizens), which they should be accountable to.

### VOPEs’ Involvement in ECD Programming

Some timid efforts were certainly made in the 1990s to involve civil society in ECD programming over the years (Waddell, 2005). The World Bank and the African Development Bank, for instance, invited the coordinators of numerous evaluation associations and leading evaluation practitioners from several developing countries to a two-day regional evaluation conference organized in Abidjan in 1999. More recently, several national governments (e.g., South Africa) expressed increasing interest in VOPE as being the entities within countries that are the best placed to enhance evaluators’ professionalization<sup>45</sup><sup>46</sup>. Despite the prominence assigned by this study to VOPE, these were not the only entities interacting with national governments on evaluation-related issues. Universities and research groups as well as NGO and community and consumer groups, for example, seemed to have

---

<sup>44</sup> Other economists (Stiglitz, 1986 1989) have long argued over the critical contribution provided by public sector’s managers to the development processes of their respective countries. However, in doing so, such authors have also recognized the distortive effects of information asymmetry on economic development and, therefore, have called for ECD programs that could enhance both the public sector managers’ demand for and access to better quality information on what works and what does not. That notwithstanding, the extent to which civil society organizations and in-country practitioners could contribute to reducing such information asymmetry was not explored at it could have, thus failing to include VOPE in any ECD-relate discourse until the start of the new millennium.

<sup>45</sup> Given that relatively large body of knowledge and practice accumulated in the area of Capacity Development (as opposed to Capacity Development specifically applied to evaluation), some authors see Capacity Development and Evaluation as two different professions (Heider, 2011).

<sup>46</sup>Despite the professionalization being not always a priority objective for many VOPE, some other authors (Morgan, 2006) point to the need for professionalizing evaluation and ECD given neither the lack of specific courses nor the lack of coherence among professionals in this area of world.

played a critical role in the promotion of an evaluative culture in a variety of countries. In particular, in some literature their “anchoring” role (Boyle & Lemaire, 1999, p.15) within the scope of evaluation institutionalization processes was recognized. Generally referred to as outside government anchors, the merit of such ECD stakeholders was to evaluate either the effectiveness of government programs or the benefits that they or their client members would derive from public services. That notwithstanding, the vital role of nongovernmental entities in the institutionalization of the evaluation function was only acknowledged by some scattered literature in 1990s. Although marginal, the contribution of skilled professionals to the strengthening of the supply side (Boyle et al., 1999) of what was referred to in this study as the ECD equation (discussed in the next section) became increasingly apparent. With such awareness, four main strategies that ought to be pursued in order to enhance such supply side were identified (Boyle et al., 1999).

First, to prepare the ground for evaluators, such as ensuring that various disciplines (sociology, psychology, political science and public administration) and competencies from which the “curriculum” for evaluators is derived are in place (Derlien, 1990) or that evaluation become a separate, standard part of academic curricula. Second, to determine the background and type of experience suitable to pursue the profession; to decide on whether one should rely managers-turned-into-evaluators, external evaluators, specialist evaluators in dedicated posts; to develop evaluators through short-term training courses, communities of practice, professional associations. Third, to develop evaluation users encouraging an active demand to match evaluation supply, that is, to enlarge the group of educated consumers (Morris, 1994). Fourth, to professionalize evaluation through the strengthening of professional associations (e.g., by making career opportunities in evaluation stable, enhancing

the certification or licensing of evaluators and disseminating standards for the practice of evaluation).

Later authors seemed to identify two additional strategies. One consisted of supporting evaluation capacity suppliers, that is, the Local Evaluation Capacity Developers (LECD)<sup>47</sup>. The other one consisted of conducting evaluations<sup>48</sup>.

Unfortunately, all such efforts and rather forward thinking strategies aimed at VOPE did not materialize as much as one would have hoped for. As a result, the degree of ownership of the evaluation function among non-governmental entities, especially in sub-Saharan Africa, was still rather low, as attested by recent online discussions among African evaluation practitioners on the AfrEA website.<sup>49</sup>

In retrospect, such rather limited, although increasing, support of civil society as part of ECD programming, did not come as surprise as the role of civil society, only marginally discussed in literature, was not be as well conceptualized, understood or disseminated as it could have. A clear illustration of this was provided by the World Bank. On the one hand, the definition of governance developed by several of the organization's economists in Washington DC, suggested that the participation of a "strong civil society in public affairs" was key to the set up of a "predictable, open and enlightened policy-making" (World Bank,

---

<sup>47</sup> Local Capacity Developers (LCD) could have ECD a core activity or just as a segment of their work. They might also be formally affiliated with an international consulting firm or an international NGO (INGO) from either the North or the South LCDs) (Ubels, 2010). LCDs include semi-public entities, such as: a) Training and research institutes; b) Leading non-governmental organizations (NGOs), c) Consultancy firms and Independent consultants South-South cooperation).

<sup>48</sup> The Development evaluation gap could be filled by systematic assessment of the whole of government policies on horizontal basis, including the involvement of VOPE This is especially true today since "the development evaluation ideas with the most traction (RBM, experimental methods, and so forth) do not emphasize the distinctive accountabilities of partners in shaping global development outcomes." (Picciotto, 2011a, p. 255).

<sup>49</sup> Such exchange dates back to September 25, 2012 and could be found on the archive of the AfrEA yahoo group website ([afrea@yahooogroup.com](mailto:afrea@yahooogroup.com)).

1994, p. 9)<sup>50</sup>. On the other hand, such a definition was never operationalized, thus failed to provide a clear framework for national governments in countries to interact with civil society organizations more effectively towards the promotion and strengthening of governance (UNEG, 2012).

A second reason for explaining such secondary role assigned to civil society in ECD programming over the years was the development partners' impossibility to support informal networks and associations, primarily for accountability reasons (e.g., their mandate to work exclusively with governmental entities).

A third reason underlying such biased targeting in ECD programming was more ideological. As CSO tended to ask questions on effectiveness and impact, that is, political questions likely to question the status quo and therefore the tenure of the government in place, development partners preferred supporting governments for the sake of maintaining peaceful relationships with their in-country official counterparts (Mayne, 2008).

Only recently, it appeared that the strive to support VOPE (a central tenet of foundations' work for over a decade), started gaining popularity among multi-lateral agencies, as attested by some of the new evaluation policies and corresponding implementation guidelines at USAID. According to its new evaluation policy, USAID started placing a special emphasis within its sectorial programming, on the support of capacity development among partner governments and civil society capacity to both undertake evaluations and use the corresponding results. In operationalizing this policy requirement, USAID also provided support to local evaluation professionals, through institutional leadership, management systems, and personnel skills improvement. The USAID evaluation

---

<sup>50</sup> Similarly, in the words of the World Bank Task Force on Capacity Development in Africa, the capacity challenge was described fundamentally as a governance challenge.

policy also stated that “to the extent possible, evaluation specialists, with appropriate expertise from partner countries, but not involved in project implementation, [would] lead and/or be included in evaluation teams.” (USAID, 2011, p. 6).

Similarly, the Director of the Independent Evaluation Group (IEG) at the World Bank, Heider, acknowledged the role of evaluation professional associations and networks and stressed how critical such entities are within national systems, primarily when it comes to supporting an enabling environment characterized by independence (structural independence), transparency and credibility (Heider 2011). In particular, she pointed to the standards set by such associations as an enhancer of evaluation professionalism but did not contemplate what role they could play in the absence of such formalization or certification. According to Heider, although VOPE were often considered an integral component of the enabling environment, evaluation associations and networks were also to be considered an integral component of both the individual level (associations members are trainees and trainers) and institutional level-interventions (associations could participate in evaluation peer-review processes and could enhance the vitality of check-and-balances and safeguards).

The tension existing within the evaluation community between the feasibility and utility of supporting VOPE as part of national evaluation institutionalization processes was reiterated. There also appeared to be a need for a more real rather than ceremonial involvement of civil society (e.g., VOPE) in the development of the national development agenda (including ECD programming). That was certainly not a new argument. As already mentioned, foundations had been supporting the strengthening of CSO for over a decade and the Joint World Bank and OECD consultations held with African leaders in September 2005 and March 2006 had already pointed to a necessary readjustment of the ECD paradigm (World Bank, 2006).

That said, according to this literature review, it could also be inferred that the lack of robust data on the effective VOPEs' contribution to the evaluation agenda in international development contexts<sup>51</sup>, prevented any specific debate on this topic from developing further.

#### Evaluation's Role in Addressing the Biased Targeting of ECD Programming

Although the VOPEs role was conceptually articulated over the years, their direct involvement in national evaluation processes could not be initiated without a proper assessment of their respective strengths and weaknesses. Without a proper understanding of VOPEs specific forces, it was hard first to figure out the type of role (including, the one of provider of evaluation services) that they could play. Similarly, without an accurate assessment of their weaknesses, it was difficult to plan ECD interventions that might be able to address the identified gaps.

As a standardized VOPE capacity needs assessment was not found before the start of this study, an instrument developed by two European Research Centers in the mid- and late 2000s was adapted to the VOPE context and used for the purpose of this dissertation. This instrument was based on two popular Capacity Development (CD) frameworks. According to the first one, developed by a South African NGO (Kaplan, 1999), capacity was “invisible” and “intangible” and it could be categorized based on five key dimensions. The second framework (Ubels et al., 2010), developed by the European Center For Development Policy Management (ECDPM) was the foundation of the Core Capability

---

<sup>51</sup> Recognizing the lack of rigorous evaluation of civil society contributions to national development programs, some efforts were made recently to fill the gap. Both the joint World Bank-DfID evaluation of the community response to HIV and AIDS in seven countries (Rodríguez-García et al., 2011) or a meta-case study analysis of a non-randomized sample of 100 research studies on citizen engagement (Gaventa & Barrett, 2010), highlighted the positive outcomes that local associations more than any other entities outside the governments) have produced in 20 countries.



Assessment, one of the most powerful tools currently available to assess changes in CD at the organizational level<sup>52</sup>. The five capabilities, although not originally conceptualized for the evaluation field but considered by this study as applicable to VOPE, were grouped in the five following categories<sup>53</sup>: a) Capability to Commit and Act; Capability to Generate Development Results; c) Capability to Relate; d) Capability to Adapt; e) Capability to Integrate. The tool aimed at assessing VOPE capacity was developed and tested in the three countries where the field data collection was conducted (Chapter III).

### Main ECB and ECD Frameworks

The tension over the use of either ECB or ECD in international development contexts as well as the asymmetries characterizing ECD targeting in the past were already introduced in the earlier two sections. In order to move toward a renewed understanding of the current debate over ECB and ECD, the main conceptual and theoretical frameworks that influenced the exchanges on this topic within the international development community over the last two decades were identified.

For the sake of clarification, given the paucity of peer-reviewed literature on ECD and the larger availability of literature on ECB, and even more so on CB and CD<sup>54</sup>, the review of the literature on ECD frameworks introduced in this chapter will borrow extensively from disciplinary areas outside of evaluation.

Far from being merely an opportunistic strategy to fill the gaps in the rather scant ECD literature, the inclusion of CD –rather than ECD-specific- frameworks and their

---

<sup>52</sup> Such tool was tested during the development of 16 global ECD case studies and was revised based on the feedback provided by those who administered in each country.

<sup>53</sup> For a more detailed description of each capability, see Appendix C.

<sup>54</sup> For the purpose of this study, CB and CD are viewed as cross-cutting themes applicable to a variety of fields including -but not confined to- evaluation

adaptation to the evaluation field, are likely to pollinate the ECD debate and possibly let some new ideas circulate among those scholars and practitioners who actively contribute to the ongoing ECD debate. With such awareness, this section aims at doing justice to the tireless efforts of those who worked on either CB, CD and, to a certain extent, ECB in the past, that is, at a time when the development of a more inclusive and participatory conceptualization of ECD is still at an incipient stage.

The three most popular frameworks dominating the discourse on ECB and ECD over the last decade and presented in this chapter were:

1. The Evaluation Capacity Development Framework of Action (Boyle & Lemaire, 1999; Boyle, Lemaire and Guerrero, 1999);
2. The Multidisciplinary Evaluation Capacity Model (Preskill & Boyle, 2008);
3. The Conceptual Framework for Developing Evaluation Capacities (Heider, 2011); and
4. The Vygotskian ECB Model (Higa & Brandon, 2008)

#### The ECB Supply-Demand Equation and the Evaluation Capacity Development Framework of Action

The first of the three frameworks presented in this chapter, and consisting of two differed tools (the ECB Supply-Demand Equation and The Evaluation Capacity Development Framework of Action) was developed in the late 1990s (Boyle & Lemaire, 1999; Boyle, Lemaire and Guerrero, 1999). The merit of this Framework was to make the ECB discourse accessible to a large audience within the international development community. In its first articulation (Table 3), the relationship existing between “donors

countries” and “aid recipient countries”(Boyle & Lemaire, 1999, p.12) was framed in terms of supply and demand.

Table 3

*The ECB Equation*

<b>Donors</b>	<b>Aid Recipients</b>
Supply of funds, training and instructions on how to conduct evaluation	Demand for evaluation services

Through the two terms (supply and demand)<sup>55</sup>, a simple and quite intuitive description of ECB activities and processes was provided at a time when the meaning of the term had not been formally defined yet.

The success of the Framework over the years was attested by the fact that a large number of practitioners framed their ECB interventions in terms of demand and supply since its release in the late 1990s. The Framework’s main assumption, which influenced the ECB planning among development partners and national governments for over a decade, was that:

- a) Donors would be responsible for the supply of:
  - Funds “to employ specialists within the governments” (Boyle & Lemaire, 1999, p.138),
  - Trainings to “enhance evaluation skills at the management level”(Boyle & Lemaire, 1999, p.138), and
  - “Instructions on how to outsource evaluation” (Boyle & Lemaire, 1999, p.138);
  - and
- b) Recipient countries would be responsible for initiating the demand for those very

---

<sup>55</sup> Both Boyle and Lemaire, consistent with their academic and professional background, borrowed such terminology from the field of economics.

same evaluation services provided by their development partners (Boyle & Lemaire, 1999).

A second assumption underlying the Framework (Boyle & Lemaire, 1999) was that, for any ECB program to be sustainable, four key elements needed to be available: (a) Financial resources; (b) Training; (c) Information; and (d) Methods and skills.

The legacy of this Framework remained strong over the years, as attested by the fact that Boyle and Lemarie's work was described in the first chapter of a recent World Bank publication focused on ECB (Rist et al., 2011)<sup>56</sup> as the conceptual foundation of numerous ECB-related debates taking place within the international development evaluation community in the early 2010s. That notwithstanding, the ECB equation (as the framework is referred to in this study) did not remain unchanged and was specified further by a World Bank ECD specialist (Guerrero, 1999). According to Guerrero's model, four different groups of ECD strategies were needed based on the different combinations of supply of and demand for evaluation:

1. First Combination (strong supply and strong demand): support to the development and implementation of national evaluation policies, strengthening of evaluation in the legislature, dissemination of evaluation results in the public, organization and systematization of the evaluation function; support to the development and management of financial and information systems;
2. Second combination (strong supply and weak demand): dissemination of evaluation methods and practices; support to ongoing evaluation of programs and projects; promotion of national evaluators' participation in evaluations done by external

---

<sup>56</sup> This publication summarizes the findings of the IDEAS Biannual Conference held in Johannesburg in April 2009

funding agencies; support of professional development in evaluation; support of research institutions in carrying out evaluations;

3. Third combination (weak supply and strong demand): dissemination of lessons learned and good practices; set up of commissions to evaluate important projects and programs; training and using of private sector institutions in evaluation, provision of technical assistance to government agencies; creation of evaluation networks within government; promotion of evaluation capacity by non-government agencies;
4. Fourth combination (weak supply and weak supply): strengthening of audit and accounting; conduct of joint evaluations with funding agencies; dissemination of national and international lessons learned; support to evaluation trainings in educational institutions; conduct of country institutional and evaluation capacity assessments, promotion of cross-country cooperation in evaluation; awareness-raising among decision makers.

On the one hand, the relevance of such variables as leadership, incentives and institutional development (Boyle et al., 1999; Dabelestein, 2003, Kuzek & Rist, 2004; Mackay, 1999; Maher, 1981; McDonald et al., 2003) was factored in by any debate on the demand for evaluation capacity, On the other hand, the debates more closely related to the supply of evaluation capacity increasingly focused on a host of new issues<sup>57</sup>, such as staffing

---

<sup>57</sup> Interestingly enough, it became increasingly apparent in the past that, while financial resources and staffing were chronically limited, the methodological and skills areas (that is, the identification of capacity needs and area of capacitation) were the most problematic ones (Bemelmans-Videc, 1992).

and skills development, financial resources and evaluation methods (Guerrero, 1999; Milstein et al., 2002).

### Critique of the Framework

Despite the Framework's popularity within the international development community, a number of limitations inherent to some of its premises were identified. Highlighting the Framework's limitations was especially relevant as it allowed gauging the extent to which the gaps associated with it were addressed by later ECB models, as those presented in the rest of the chapter. Overall, three main limitations were identified in relation to the ECB Supply-Demand Equation and the Capacity Development Framework of Action.

First, through the use of such generic terms as "recipient country" associated with the demand side of the ECB equation and "donors" as the corresponding supply side, the diversity of actors associated with each side of the ECB equation, and operating either within or outside the national government's sphere, was not captured. Second, the rigid classification of recipient countries as "demand agents" and of donors as "ECB suppliers," did overlook the fact that entities associated with recipient countries and donors could play an effective role on both sides of the equation and that the relations (and interactions) between the "demand" and the "supply" were much more fluid than initially envisaged (Nielsen and Attström, 2011). As a result, the fact that VOPE may serve not only as recipients but also as providers of evaluation services, was not adequately acknowledged by the Framework. Third, despite the practical and rather generic recommendations provided for each of the four possible scenarios (Strong Demand-Weak Supply, Strong Demand-Strong Supply; Weak Demand-Weak Supply; Weak Demand, Strong Supply), two main deficiencies were identified in association with the Model. On the one hand, not too much

guidance was provided on how to determine the strength level of either the supply of or the demand for evaluation. On the other hand, as reality is seldom either black or white and, therefore, the ECB demand and supply might be neither strong nor weak but rather be situated in the middle, the broader variety of existing possible capacity needs scenarios was not captured effectively by the Framework.

### The Multidisciplinary ECB Model

The second ECB theoretical frameworks that was identified and analyzed in this chapter was the Multidisciplinary ECB Model developed by two scholars based in the U.S. (Preskill & Boyle, 2008). This Model, which was disseminated nearly a year after the release of the “ECB supply and demand” Framework discussed in the earlier section, appeared conceptually very appealing and reflected a solid understanding not only of ECB strategies but also of how such strategies fit within the organizational project cycle (Figure 1).

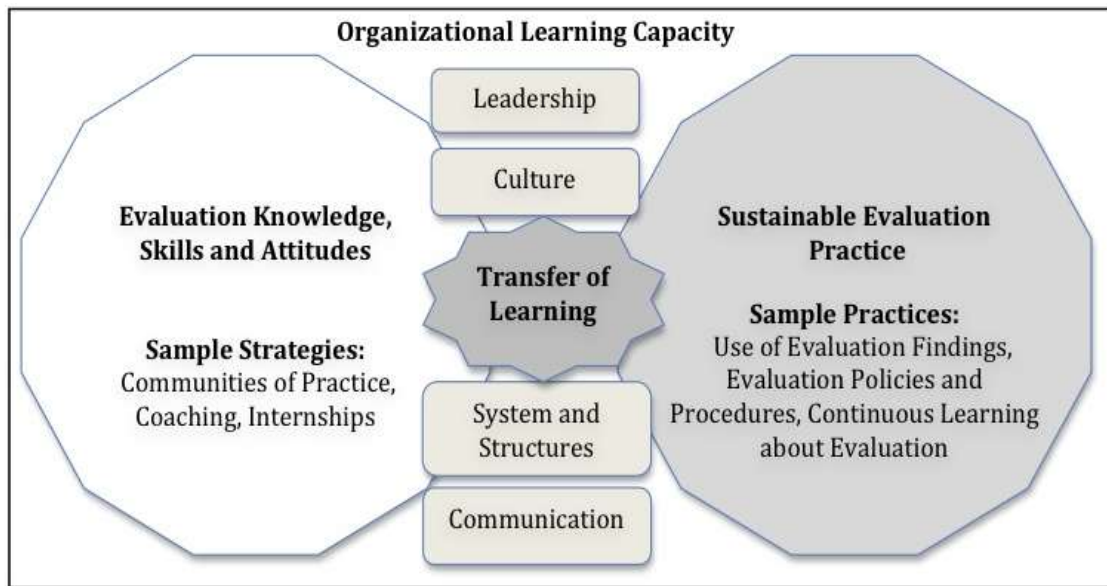


Figure 1. A Multidisciplinary Model of Evaluation Capacity Building (ECB)

The Model's main contribution consisted in having introduced a more system-oriented and realistic description and interpretation of ECB that the Framework presented in the earlier section has failed to capture. The model was more systemic-oriented<sup>58</sup> than the framework presented in the earlier section for two main reasons. First, the interdependence existing between “evaluation knowledge/skills/attitudes” and “sustainable evaluation practice” (large circle on the right) was amply recognized. Second, the influence exerted on the organizational learning capacity<sup>59</sup> by communication, systems & structures as well as culture and leadership; was stressed. Overall, the model set itself apart from other frameworks that had instead stressed the relevance of only a few of ECB foundational components or elements, such as financial resources, information, location and values (Morgan 2006).

In addition, the Model was more relational and political than the prior Framework in that it acknowledged that both the supply of and demand for ECB-related services (e.g., the request for more ECB made by citizens, clients, politicians and stakeholders) were profoundly affected by a large variety of ecosystem factors (Boesen & Therkildsen, 2005; Kuzmin, 2004), comparable to forces (either external or internal) with either a political or functional role (Table 4).

---

<sup>58</sup> Given the increasing need for evaluating more systemic programs, more ECD evaluation will need to take place in the future. As Picciotto states: “The privileged units of account of development evaluation are individual projects or country programs. Evaluation of global policies and collaborative initiatives that shape the international response to global crises will need increasing attention” (Picciotto, 2011, p.254).

<sup>59</sup> This feature of the Model resonates quite well with some of the most recent ECB and ECD literature that both highlights the importance of relationship capacities (Crutchfield & Grant, 2008; Bloom & Dees, 2007; Morgan, 2006) and stresses the connection existing between any given organization – an open system by nature– and the larger ecosystem where they operate. As two ECB researchers once stated: “*High impact non-profit work with and through other organizations – and they have much more impact than if they acted alone.*” (Crutchfield & Grant, 2008, p. 21).



Table 4

*The Four Forces Influencing ECB Capacity (By Level and Role)*

	<b>FUNCTIONAL ROLE</b>	<b>POLITICAL ROLE</b>
<b>ORGANIZATIONAL LEVEL</b>	<ul style="list-style-type: none"> <li>• Organization’s mission and values, organizational evaluation policies and approaches</li> <li>• Definition of roles and responsibilities among evaluation stakeholders, establishment of rules guiding the evaluation reporting processes, etc.</li> </ul>	<ul style="list-style-type: none"> <li>• Evaluation champions’ leadership,</li> <li>• Distribution of power between evaluation planners and decision-makers</li> <li>• Monetary and non monetary incentives for the sound implementation of the evaluation function ,</li> <li>• Conflicts of Interest</li> </ul>
<b>INSTITUTIONAL LEVEL</b>	<ul style="list-style-type: none"> <li>• National Evaluation Policies</li> <li>• International Treaties providing for ECD and the use of evaluation, timeliness and adequacy of resources to conduct evaluation,</li> <li>• Performance targets,</li> <li>• Oversight bodies,</li> <li>• Formal accountability requirements</li> </ul>	<ul style="list-style-type: none"> <li>• Political governance,</li> <li>• Third-party interests,</li> <li>• Pressure from clients, customers, competitors</li> <li>• Media attention</li> </ul>

Overall, the merit of this Model was to have emphasized the role of the ECB processes, rather than its activities. In this sense, the Model was built upon the concept of relationships, described as a critical component of any capacity development process (Bloom & Dees, 2007; Morgan, 2006). Although the complexity of the system where ECB took place was not overemphasized by the Model, the connection between the individual organizations –regarded as an open system – that demand and/or provide ECB; and the larger ecosystem with its complexity of processes, opportunities and challenges (Crutchfield & Grant, 2008), where these organizations operate<sup>60</sup>, was clarified.

---

<sup>60</sup> The interconnectedness highlighted in this Model clashes with some of the ECD frameworks presented in the next section. Such frameworks tend to be generally more structured around three different levels (individual, organization and enabling environment).

In contrast with some other ECD frameworks (such as the one introduced in the next section) that dissected ECB and ECD discourse across three different levels (individual, organizational and institutional), such levels were not depicted by the Multidisciplinary Model as stand-alone blocks with a life of their own, but were described and interpreted rather as if they were in continued interaction with each other.

This model was also more realistic than the “supply-demand” framework presented earlier, in that it went beyond a merely project-based ECB approach (more limited in scope and duration) and called for: a) more sustainable evaluation practices (Schröter, 2008), such as the provision of sufficient resources for conducting evaluation; b) a more systematic use of evaluation findings; c) the sharing of common values and beliefs among evaluation practitioners (Joint Committee on Standards for Educational Evaluation, 1994); and d) the creation of evaluation procedures and policies harmonizing ECB efforts at the national level.

#### Critique of the Model

Despite the Multidisciplinary Model’s contribution to the current ECB discourse, two main weaknesses were associated with it, when applied to the international development evaluation context. First, national evaluation associations and Communities of Practice (CoP) (Wenger, 1998) were described in the Model as simple ECB “initiatives,” “activities” or “strategies” (depicted at the bottom of the left circle) and were not assigned, as argued in this study, a more prominent role. As the CoP or VOPE were considered in this study as critical “agents” of change enhancing the sustainability of evaluation endeavors at the national level, they should have been placed, along with the other influencing or “ecosystem factors,” in the area between the circles, if not in a third circle altogether.

Second, the Preskill and Boyle Model was not able to account for the complexity of evaluation systems. Despite being linked to the systems thinking paradigm and recognizing the interdependence among ECB stakeholders (as well as the influence of systemic forces on them), the model did not appear sufficiently dynamic. Some of the factors influencing the successful implementation of ECB (the motivation for demanding and supplying ECB; the assumptions and expectations about ECB; and the ECB goals and objectives), for instance, were not included in the model and most of the learning strategies and sustainable practices were described as isolated (rather than interacting and evolving) elements of a system. As such, a multi-level, rather than cyclical, representation of ECB process would have been more accurate.

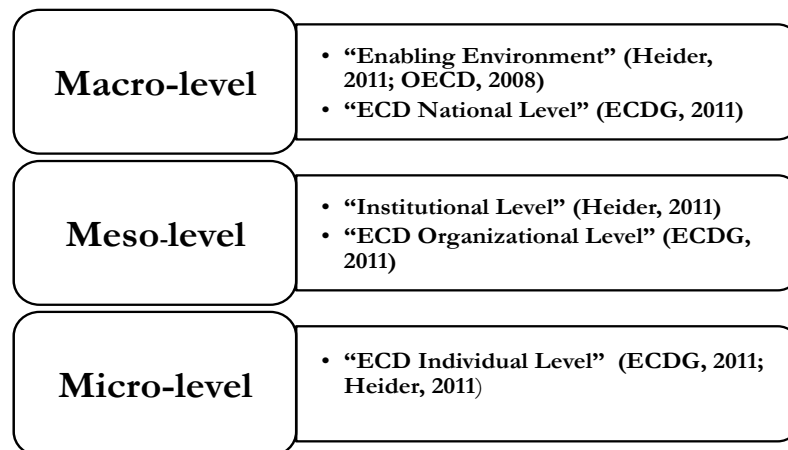
#### The Three-level ECD Framework

It was only recently that the concept of ECD assumed the connotations of a three-level construct, as attested by a recent work presented by Heider (currently the Director of the Evaluation Group at the World Bank) at the Biannual IDEAS conference in Johannesburg in 2009. This third Framework, known as the *Conceptual Framework for Developing Evaluation Capacities*, could be more simply referred to as the three Level ECD Model. The model's main contribution was to contribute to a shift in ECD conceptualization from a shortsighted strategy consisting of individual trainings to a better-articulated intervention delivered at three levels:

1. Individual level: activities enhancing knowledge, skills and competencies (Taut, 2007);

2. Institutional or Organizational level: activities enhancing a system and structure in which individuals can perform and attain results individually as well as collectively as an organization (King & Volkov, 2005; Preskill & Boyle; 2008); and
3. Environmental level: activities fostering the performance and results of individuals and organizations (Furubo, Rist & Sandahl, 2002; Heider, 2011; Mackay, 1999).

This third Conceptual Framework represented a systematic effort to a) rationalize the web of ECD stakeholders at different levels (individual, organizations and institutions), and b) organize their respective processes and influencing factors in hierarchical terms (micro, meso and macro level) (Heider 2011). Although it was consistent with the same system perspective, which the Preskill and Boyle Model was founded on, this Conceptual Framework (Figure 2) appeared to be simpler and more intuitive than the Multidisciplinary ECB Model



*Figure 2.* The Three Level Model: Key Elements and Major Contributing Authors

This Conceptual Framework represented an enhanced version of the CIDA *Conceptual Framework for Capacity Development* (CIDA, 2000) developed in the late 1990s and articulated across four different levels of capacity (OECD, 1995; UNDP, 1998)<sup>61</sup>: (a) individual, (b) organizational, (c) network/sectorial and (d) the enabling environment.

### Critique of the Framework

Despite the Model's positive contribution to the ECD discourse, three main limitations were associated with it. First, by placing the enabling environment at the macro level (the highest of the three hierarchical levels), the Model failed to take into account that the two other levels (the individual and organizational/institutional levels) were also affected by what might be considered environmental factors at their respective (rather than higher order) level. This was the case, for instance, with the lack of internet access in rural areas which placed huge constraints on individual evaluation professionals' access to online training, despite their interest in enhancing their quantitative analysis skills or their surveying techniques. Likewise, this was the case of those evaluation consultants who, due the economic crisis affecting their communities and the inevitable repercussions on their own personal finances (e.g., less assignment opportunities available to them), were more easily inclined than usual to review the negative findings of their evaluation in favor of their clients so as to get more contracts from them in the future.

Second, by dismissing the systemic features of the Preskill Model, the Three Level Framework, though simpler and more intuitive than the prior, was responsible for perpetuating a certain fragmentation of the ECD discourse (ECDG, 2012). In doing so, it

---

<sup>61</sup> Each of these represents a level of analysis, as well as a possible entry level for an ECD intervention.

seemed to have missed the opportunity to a) make the connection among the three levels more explicit and b) avert the ECD stakeholders' common practice of focusing either on one level only (individuals) or, more rarely, on two levels (individual and organizational).<sup>62</sup> As a result, the framework overlooked the fact that ECD was very political and relational: its activities and processes were not isolated and stand-alone. The interactions among the three levels were critical to the success of ECD in any given country.

Third, what was also missing in the Model was the concept of "citizen's control," that is, a group of individuals politically organized more or less consciously and not necessarily associated with a given program, project or organization.

#### The Vygotskian ECB Model

Through the Vygotskian ECB Model (Brandon & Higa, 2004) the centrality of learning within ECB processes was recognized and a particular emphasis was placed over the relevance of contextual variables e.g., type and quality of social interactions among ECB stakeholders) as mediating factors of learning. Echoing the work of Vygotsky, father of Social Development Theory (McLeod, 2007; Vygotsky, 1997), this model stressed that the success of any ECB endeavor was affected not only by the number of people involved in evaluation (the so-called critical mass) but also by the quality of ECB content. According to this model, while individuals may develop a more positive attitude to evaluation and accordingly develop some technical skills to both conduct and use evaluation more consistently as a result of their participation in one or more ECB activities, it was only

---

<sup>62</sup> UNDP and UNICEF have focused on both the meso- and the macro-level. Some other organizations, such as ECDG, ECDPM and other "enlightened" evaluation training institutes have tried to target entire organizations rather than individual members.

through the engagement in evaluation of a large number of actors within an organization that an evaluation organizational memory could develop.

According to Higa and Brandon (2008), the number of people involved in evaluation and the quality of their involvement were directly related to the quality of the evaluation activities. These authors proposed that more involvement generally resulted in a longer organizational memory, as long as the involvement of each individual was meaningful and enabled them to build their own personal evaluation capacity.

#### Critique of the Model

Although the application of a psychology theory to evaluation was a commendable effort, the focus on the ecological aspect of ECB predicated by this model risked being reductive in that it did not adequately capture the more subjective (rather than structural) aspects of the interactions amongst stakeholders and between stakeholders and the surrounding environment. In addition, as recognized by the same authors of the research on the use of such model in the evaluation of a specific ECB program (Higa & Brandon, 2008), the high implementation costs of conducting interviews and interventions over a longer period of time to test its validity, proved to be a barrier to its replicability in future research.

#### The Need for a New ECB/ECD Framework

In order to address the weaknesses of the four prior frameworks and based on the findings of the systematic review undertaken during the first phase of data collection, a new ECB and ECD framework was developed and presented in Chapter IV.

This framework, validated by a number of evaluation practitioners during the first two phases of data collection, was used to assess the interaction among ECB/ECD stakeholders in the three countries where the case studies were conducted.

## ECD Evaluative Frameworks

### Introduction

Given both the large variety and systemic objectives of activities implemented as part of ECD programs<sup>63</sup>, evaluating their effectiveness in a comprehensive manner has been quite challenging to date. The difficulty of assessing the performance of ECD programs is even more apparent in one factor in the long-term nature of most ECD objectives (e.g., institutionalization of the evaluation function within the government or the mainstreaming of more rigorous evaluation practices within VOPE)<sup>64</sup>(Wing, 2004):

“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, p. 10), and “are not necessarily attributable to specific CD processes or approaches” (James, 2001, p. 8).

In order to enhance the understanding of the current state of ECD evaluation practices and as a way to dissipate some of the fears surrounding their implementation, this section reviews four of the major CB and CD evaluative frameworks used within a variety of international development agencies.

### Evaluative Frameworks and Approaches

#### First Framework: Kirkpatrick Four Level Evaluation Model

The first framework, known as the Kirkpatrick Four Level Model was developed in 1998 and over the years it has become the evaluative framework most widely used to assess the effectiveness of trainings both at the individual and organizational level (Table 5).

---

<sup>63</sup> As Morgan states, “capacity development becomes almost synonymous with development itself, making it difficult to assess results in a systematic way” (Morgan, 1998, p. 2).

<sup>64</sup> The long timeframe deemed necessary for ECD interventions to demonstrate results entails the high risk of drop-out among individuals included in treatment and comparison groups as well as the difficulty of isolating ECD activities as the only causal factors of change.



Table 5

*Kirkpatrick Four Level Evaluation Model*

Level	Definition	Measuring tools
<b>Level 1</b>	Participant satisfaction	End-of-course participant questionnaires.
<b>Level 2</b>	Learning outputs	Posttests, sometimes as compared with pretests.
<b>Level 3</b>	Performance change outcomes	Multiple, including observation, interviews and surveys of participants, colleagues, and supervisors.
<b>Level 4</b>	Organizational impact/results	Multiple, including comparisons with baseline organizational performance measures, surveys, and interviews with key informants

Source: Kirkpatrick (1998)

Despite its clarity and popularity across a variety of sectors, this model was not always utilized in an orthodox fashion over the last 20 years. In particular, through a review of the existing evaluation literature, it was concluded that a large gap existed between the frequency of measurement of the first two levels (Levels 1 and 2) and that of the last two levels (Levels 3 and 4).

On the one hand, the results defined as Level 1 (participant satisfaction) and Level 2 (learning outputs) are measured consistently across interventions. The relative ease in the collection of and access to such data was confirmed by literature in numerous cases. A recent evaluation conducted by the WB Independent Evaluation Group (IEG) and aimed at assessing the effectiveness of the World Bank-funded trainings delivered between 2000 and 2005 across a variety of sectors, attested great results in terms of individual participant learning (Kirkpatrick Level 2) (World Bank, 2008). Similarly, in a recent systemic review of ECB interventions (Labin et al., 2011) the most frequent outcomes reported in 92% (57/62) of the studies were individual-level outcomes. The specific outcomes that were most

frequently measured were related to changes in attitudes and knowledge, as found in 51% (32/62) of the total studies included in the systematic review.

On the other hand, the measurement of Level 3 (performance change outcomes) and Level 4 (organizational impact) results was much more challenging than for the previous two levels. That such variables were harder to both attain and measure did not come as a surprise, as reiterated by the same WB IEG report mentioned earlier, according to which the development of organizations' or trainees' capacity to achieve development objectives (Kirkpatrick levels 3 and 4) occurred only in half of the studies<sup>65</sup> included in its analysis (World Bank, 2008).

Likewise, in a recent report commissioned by UNESCO (Ortiz & Taylor, 2008) the challenges of assessing the organizational level outcomes (Level 4 result) of ECD interventions was reiterated:

Organizational priorities – defined emergently – [*seem to*] drive the process, resulting in customized, self-led [E]CD (allowing the organization time to accompany its own process). [At this point] 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 [E]CD interventions that they fund dedicated organizational learning time to process [E]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 [e.g., in funding use] to this type of change (Ortiz & Taylor, 2008, p. 40).

As a result, it did not come as a surprise that, according to the same systemic literature review of ECB interventions mentioned earlier (Labin et al., 2012), the number of organizational-level variables reported in 77% (48/62) of the studies was lower than the corresponding number of individual-level variables. This seemed to support the argument

---

<sup>65</sup> This was mostly due to such variables as: lack of incentives, inadequate targeting, trainees' poor understanding of how to use the knowledge or lack of customization to the trainee's needs and context

that ECB was mostly about individual-level training and not so much a systemic approach<sup>66</sup>. Overall, of all the studies reporting quantitative data, 20% (13/62) reported individual-level outcomes and only 10% (6/62) reported organizational-level outcomes (Labin et al., 2012). That notwithstanding, for the studies included in the research synthesis mentioned above (Labin et al., 2012), reporting organization-level outcomes included in the review (24/62), Processes, Policies and Practices (PPP) appeared to be the most frequent type of outcome followed by mainstreaming and resources. Culture and leadership were the least common outcomes reported in the studies included in the review. According to 54% percent (13/24) of such studies, multiple strategies were particularly effective in promoting mainstreaming, including: a) allocation of additional resources; b) increased leadership support; and c) the introduction of a new culture. Furthermore, leadership was one of the most difficult outcomes to affect and so was the case with mainstreaming. Interestingly enough, leadership was the least frequently targeted organizational factor and the last frequently reported organizational outcome. (Labin et al., 2012).

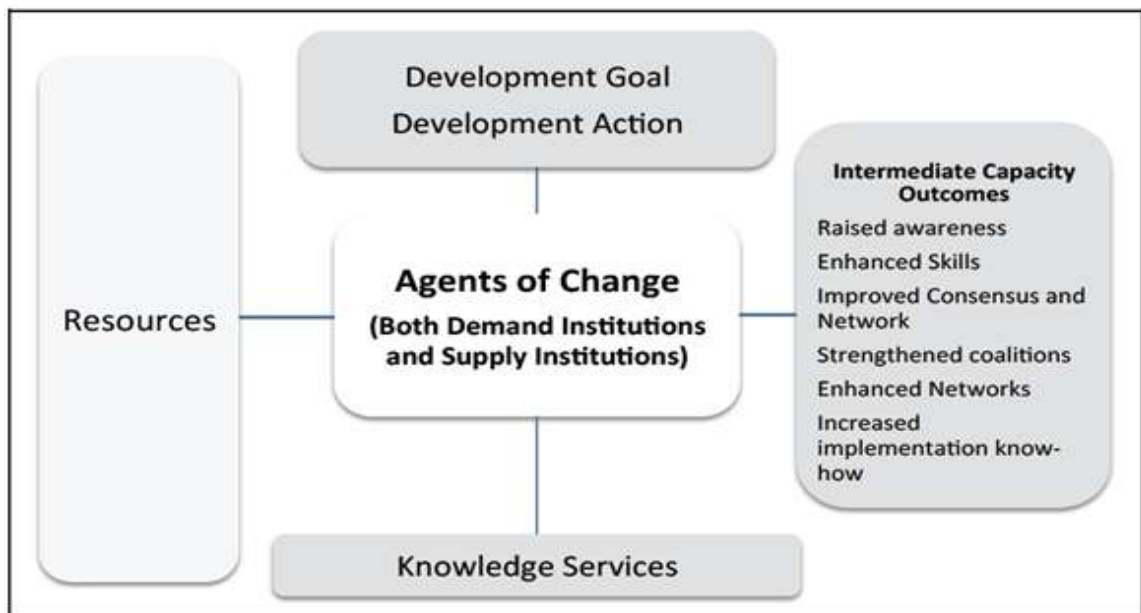
#### Second Framework: The World Bank ECB Results Framework

The second evaluative framework used to evaluate the effectiveness of evaluation capacity building was known as the World Bank ECB Results Framework (World Bank, 2011). Its merit consisted in providing a concise and yet quite comprehensive conceptual representation of ECB critical components. Two main strengths were associated with this

---

<sup>66</sup> According to the same study (Labin et al., 2012), leadership support in 32% ( $n=20$ ) as well as resources in 36% ( $n=22$ ) were mentioned as critical variables. Interestingly enough, among the support of organizational-level outcomes, the provision of incentives or further development of leadership support was attested by only 9.8% of the reports ( $n=6$ ). The lack of resources was also reported in 49% ( $n=30$ ) as the most common organizational-level barrier to evaluation. Likewise, the staff turnover was identified as one of the main barriers (23% or  $n=14$ ). Therefore, it is a necessary thing to do but it is often too expensive.

framework. First, its ability to speak to those who design and manage ECB within the World Bank by showing them what was needed in order to achieve development objectives through ECB in a more systematic fashion (Figure 3). Second, its rather comprehensive definition of ECB, no longer regarded as simply technically based and compliance driven: “development objectives” were defined by the framework as “improvements in stakeholder ownership, policy-related and organizational capacity areas” (World Bank, 2011, p. 25).



*Source:* Adapted from World Bank, 2011

*Figure 3.* ECB Results Framework

However, a few limitations were associated with the framework. Although a new comprehensive way of measuring ECB effectiveness was predicated by this model, no insight was provided on how to implement ECD in a more sustainable and participatory manner (the provision of knowledge services from the World Bank to its clients was clearly depicted by a one-way arrow pointing upward to an undefined “group of agents of change”). Likewise, the concept of supply-demand associated with the classic ECB Framework (Boyle

et al., 1999) was still evoked, thus ignoring the fluidity of roles among ECD stakeholders (Figure 4).

Raised Awareness		Enhanced Knowledge and Skills	
Improved Consensus and Teamwork	Strengthened Coalition		Enhanced Networks
New Implementation of Know-how			

*Source:* Adapted from World Bank, 2011

*Figure 4.* Six Intermediate Capacity Outcomes

### Third Framework: The Results-Based Management (RBM) Evaluative Approach

Also known as “logframe-based” evaluation approach (Coryn, Noakes, Westine, & Schröter, 2011), the RBM<sup>67</sup> Evaluative Approach, more closely linked with what this study defines as ECB, suggested that the best proxy for capacity development was the improved performance of the individuals capacitated against a clear set of predefined learning objectives.

This approach certainly had a number of merits. First, the emphasis on accountability as attested by the propensity of ECD programmers and implementers to measure program performance as a way to verify the extent to which ECD programs were delivering against the specific capacities and corresponding levels of success identified during the design phase (Baser and Morgan, 2008). Second, the focus on the attainment of pre-identified objectives (that is, the objectives found in the logical framework and/or results

---

<sup>67</sup> The term RBM is particularly popular among development professionals. Widely used in the 1990s and 2000s (some practitioners have increasingly replaced RBM with Management for Development Results since 2008), RBM rests of 4 pillars, including M&E. According to a recent UNEG definition, RB is “a management strategy by which all actors on the ground, contributing directly or indirectly to achieving a set of development results, ensure that their processes, products and services contribute to the achievement of desired results (outputs, outcomes and goals). RBM rests on clearly defined accountability for results and requires monitoring and self-assessment of progress towards results, including reporting on performance” (UNDP 2007, p.7).

strategy) and the subsequent relevance of capacity needs assessment (discussed more thoroughly in one of the following sub-sections), proved to be key to better targeting and customizing ECD interventions to both individual and institutional needs. That notwithstanding, the RBM Approach was not spared criticism for promoting what some authors considered to be a technocratic/reductionist thinking (Baser & Morgan, 2008). First, the RBM critics argued that it was critical to take the less tangible and more relational/attitudinal dimensions of capacity into consideration—and that, therefore, the planning—and control-oriented “reductionist approach” was not adequate to measure the unexpected consequences and uncontrollable causes of divergences of an ECD intervention from the agreed upon programmatic road maps. Second, the RBM evaluative Approach, also referred to as the “checklist approach,” was often criticized for turning CSO into “inert vehicles for delivering particular services or projects whose conditions had already been established by the funders“ (McAllister, 2011, p. 210). Third, the RBM Approach was often criticized for being more of a tool to test program staff’s predictive rather than adaptive capacity (see Figures 5).

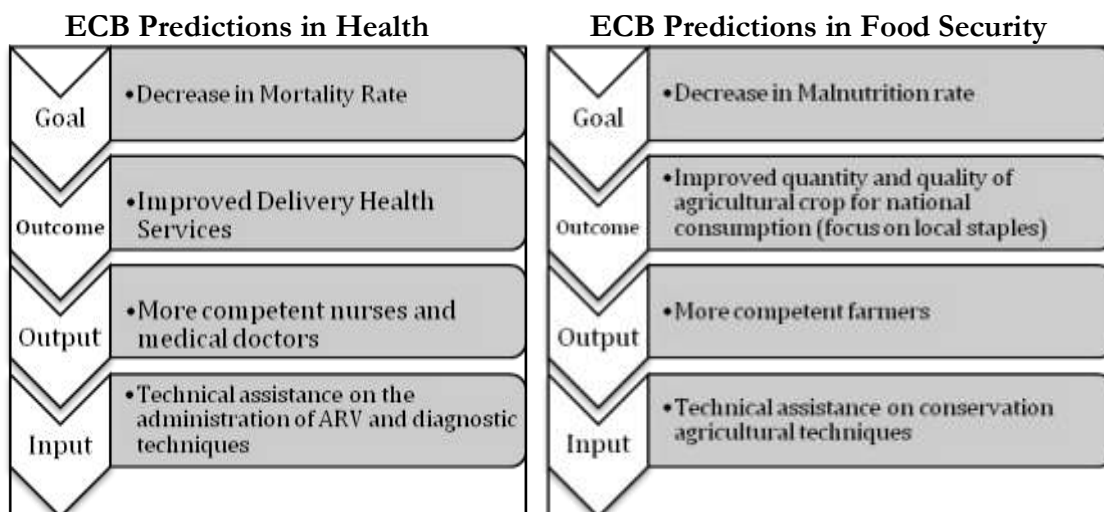


Figure 5. Examples of Predictive Capacity in ECB Planning

World Bank and USAID, increasingly subject to the scrutiny of member countries and taxpayers respectively, favored this approach more than other agencies. However, the RBM became increasingly popular with national governmental agencies in the early 2000s. Such was the case of the Rwanda Revenue Authority, the Philippines-Canada Local Government Support Programmes, and the Faisalabad City of District Government in Pakistan (Watson, 2010). Overall, this approach proved to be more effective (Watson 2010) when the individuals/institutions being capacitated did the following: a) volunteered to receive capacity development support; b) were capable of assessing the capacities they need; c) were able to clearly define what abilities they needed to acquire or improve; and d) received incentives to enhance their performance, and e) benefited from their organization's leader support.

#### *RBM and Capacity Needs Assessment*

If the success of a program was measured by its ability to meet its expected objectives, then the success of an ECD program rested on its ability to address a set of capacity-related needs<sup>68</sup>. As needs were expected to be identified during a pre-evaluation phase, normally known under the name of capacity needs assessments<sup>69</sup>, then the latter was a

---

<sup>68</sup> Interestingly enough, two recent studies using DAC bilateral data finds that almost half of the predicted value of aid is determined by donor-specific factors, one-third by needs, one-sixth by self-interest and only 2% by performance (Hoeffler & Outram, 2008; Rogerson & Strensen, 2009)

<sup>69</sup> Evaluation Capacity Needs Assessment is a front-end 3-step inquiry (pre-assessment; assessment; post-assessment) consisting in the: a) Identification of already existing information on ECD needs among local evaluation national associations – this will inform the determination of the needs assessment focus; b) Collection of new data additional to what has not been learned yet during the first phase – this will inform the in-depth investigation of the needs at three different levels (ECD program recipient, ECD program deliverers, donor community and national systems) as well as of the causes and needs priorities; and c) Prioritization and comparison of ECD alternative program and strategies for decision makers to inform the design, formulation and evaluation of ECD program in the future (Altschuld, 2010a,b,c,d,e).

necessary condition for a focused evaluation (Watkins et al., 2012). That notwithstanding, capacity needs assessments were not conducted as frequently as one might expect.

According to the systemic review of ECB interventions (Labin et al., 2012), although only 30% (18/62) of the reports being reviewed confirmed the use of a needs assessment, nearly 26% percent (16/62) did not make any reference to it, and some 64% percent (39/62) mentioned that tailoring/customization was quite common.

Concentrating on need assessment in an under-researched area, such as ECD was all the more relevant, as policy-makers and program managers often based their ECD programming decisions on assumptions and values that, although shared by the international development community, were the “same old solutions” (Watkins et al., 2012), that is, ECD initiatives were rarely informed by timely and context-specific evidence as well as responsive to actual needs. The idea that implementing a needs assessment was conducive to any development intervention (not exclusively in the area of ECD) also built on Scriven’s Evaluation Training Checklist (Scriven, 2008) in which the author stated the following:

“[...] All too often the supposed need is merely one of the following: a long-established offering, not recently reconsidered; an unsubstantiated intuition by some executive who has been inspired by something they heard or read; the results of a wants survey of staff (who may really be voting for a fun change from boring routines); something the HR department thinks would make them look with-it or at least useful; or a ‘keep up with the Jones’ response to some current fashion in training or anecdotal report from (or about) a competitor” (Scriven, 2008, p. 5)

“[...] Providing this kind of needs assessment in detail could be a major task requiring considerable skill, but it is typically much cheaper than undertaking training based on someone’s hunches about these issues” (Scriven 2008, p. 9)

In particular, what seemed to be a major gap in this area was the recurrent practice to consider needs as “needed solutions” (the recurrent saying is “we need a new policy” or “we



need a new training in mixed methods”) rather than inquiring on what the real problem is and what the “gaps in results” were <sup>70</sup>

### The Complex Adaptive Thinking (CAT) Evaluative Approach

The fourth framework, known as the Complex Adaptive Thinking (CAT) Evaluative Approach, was the most closely associated framework linked with what this study described as ECD. It was radically different from the RBM approach and some of its premises originated from the World Bank ECB Framework.

Aimed at capturing the complexity<sup>71</sup> of the context where the ECD program took place, the CAT evaluative approach was characterized by a more systemic thinking perspective of ECD<sup>72</sup> as well as a large use of interactive M&E methods (e.g., use of stories, interactions with ECD stakeholders and participatory methods)<sup>73</sup>. This approach was well grounded on the assumption that, “as capacity is associated with multiple causes, solutions and effects, some of the corresponding dynamics are not necessarily controllable and potentially quite unpredictable” (Watson, 2010, p. 26)<sup>74</sup>.

---

<sup>70</sup> The identification of the gaps in results should be followed, during data analysis, by the prioritization of needs. In order to do that, one will need to compare the costs associated with addressing each of them (or closing the gap) and (b) the costs associated with not addressing each of them (or leaving the gap).

<sup>71</sup> Complexity thinking (Ramalingam et al., 2008) is not a prerogative of the development field, As Alan Fowler wrote in an article on the failure of development aid (Fowler, 2008), he reminds the reader that Robert Axelrod and Michael Cohen of the University of Michigan had already explore ways in which organizations can ‘harness’ complexity, and that Margaret Wheatley, president emerita of the Berkana Institute, has applied complexity analysis to leadership. Complexity thinking – Fowler reminds the reader- was also used in politics and terrorism analysis.

<sup>72</sup> This perspective seems more aligned with the definition of capacity as an emergent combination of attributes, assets, capabilities and relationships (ECDPM, 2008) that enables a human system to perform, survive and self-renew.

<sup>73</sup> Complexity evaluations was described by Michael Patton as one of the 10 major trends in evaluation during an event hosted by the US State Department on July 23, 2012 (Patton, 2012).

<sup>74</sup> The concept that reality is not as static or easy to codify (e.g., the identification of clear roles and responsibilities among development program stakeholders) has also been a leitmotif of the modern management theory according to which businesses are the result of value-chain relations rather than the function of processes and tasks that can be analyzed in isolation. ECD programming could definitely benefit from the paradigm dominating the contemporary management field which mostly consists of four stages: 1) Identify bottlenecks; 2) Remove friction 3) Shrink to size; and 4) Adjust to demand

The added value of such new M&E approaches, especially if compared with donor “accounting” approaches, was clearly identified in a recent work published by the Institute of Development Studies at the University of Sussex (Ortiz and Taylor, 2008):

“Observation and study, learning, abstract framing, adaptive management and agility in changing plans and putting learning into practice are more important than rigorous tracking of outputs that ultimately do not reflect at all the reality of the situation they are describing” (Ortiz and Taylor, 2008, p. 33).

Such an approach was certainly used to assess the level of technical performance of those who had been capacitated. However, in going beyond the typical ECB evaluations, the CAT evaluative approach was aimed at assessing the degree to which the strengthening processes embedded within ECD interventions contributed to developing structural and not simply operational capacities (Kaplan, 1999; Morgan, 2006) further. Such capacities were more intangible than those measured by the enhanced knowledge and application of a certain technical procedure. The CAT approach assessed such capacities as relationship leverage, capacity needs assessment design capabilities and internal evaluation culture.

The CAT approach was embraced by a smaller group of organizations. Both the Australian Agency for International Development Agency (AusAID) and the German Federal Ministry of Economic Development Cooperation (BMZ) abandoned logical frameworks as a planning and management tool before everybody else. In both cases, only outcomes and corresponding indicators were determined beforehand. The decision on what inputs to employ, activities to implement and outputs to produce was left to the team implementing the programs in the field. As a result of this CAS Framework there was enough room to revise and change all such “fine details,” depending on the circumstances and needs.

Additional examples of successful implementation of the CAT approach were the evaluation of the Environmental Action (ENACT) programme in Jamaica or of the Asian chapter of the International Union for the Conservation of Nature (IUCN). Based on these two cases, the effectiveness of this approach was enhanced by several factors. First, the flexibility demonstrated by donors (e.g., the Canadian International Development Agency in the ENACT case) (Watson, 2011) to address quite promptly the uncertainties that they, along with the implementing agencies, faced during the implementation of the program (the goal was there but the roadmap was not set in stone). Second, the institutions being capacitated were allowed and encouraged to learn lessons from their own experiences and were able to come up with ECD strategies to pursue in the future. Third, mentoring and experimental learning<sup>75</sup> rather than training were the preferred capacity enhancing modalities (Watson, 2010). Such an approach was also heavily influenced by Complexity theory, according to which it was not possible to predict with any confidence the relation between cause and effect. According to the complexity theory, “change is emergent and history is largely unpredictable” (Ortiz & Taylor, 2008, p.27). As a result, “small ‘butterfly’ actions may have a major impact while larger interventions may have very little impact” (Eyben et al., 2008, p.203). In order to capture such an unprecedented evolving scenario<sup>76</sup>, the CAT

---

<sup>75</sup> This last feature of the approach resonates quite well with the cultural beliefs of countries with a large number of Buddhist followers (e.g. in Asia): Confucius predicated that humans may learn wisdom by three methods: 1) by reflection (the noblest of all); 2) imitation (the easiest); and experience (the bitterest).

<sup>76</sup> As Ortiz & Taylor put it: 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 (Ortiz & Taylor, 2008). No one is in overall control [*anymore*] of what is happening, and although patterns of relating tend in a particular direction, the exact global pattern that emerges is unpredictable (Mowles, Stacey & 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, p.815).

Framework envisaged of the use of iterative planning and measurement approaches as stated by Bakewell and Garbutt (2005):

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 towards the goal without stepping off a solid path and becoming possibly lost. [...] The concept here is to develop initial consensus on an intentional short time frame and then continually implementing re-planning as information and smart directions emerge (Bakewell & Garbutt, 2005, p. 20)

The evaluation methodologies that seemed to be the most promising in assessing ECD effectiveness were:

- a) The Most Significant Change technique (Dart & Davies, 2004);
- b) Accountability, Learning and Planning in Action Aid (Guijt, 2008);
- c) Outcome Mapping (Earl et al., 2001);
- d) Critical Incidence Technique (Hettlage & Steinlin, 2006);
- e) Rich Picture Drawing (Checkland, 2008);
- f) After Action Review (Serrat, 2008);
- g) Mini case studies (Flyvbjerg, 2006; Yin, 2003);
- h) Journal/diary writing

Several merits were identified in relation to such new approaches. First, ECD stakeholders were encouraged to reflect upon, appreciate as well as question the relevance, utility and effectiveness of ECD activities (e.g., the learning as well as the process in sharing

and building upon the seemingly enhanced capacity). Second, mixed-methods were heavily relied on and therefore contributed to the pursuit of exploratory and confirmatory studies. Third, a larger number of stakeholders were allowed to express their ideas (both individually and collectively) on the ECD contributions, both within and outside the organization<sup>77</sup>. Despite their added value, such methods were not yet part of the practitioners' toolbox that was most widely used in international development contexts. According to a recent systemic review of ECB interventions (Labin et al., 2012), for instance, the data collection methods most frequently used to assess the effectiveness of ECB effort continued to be a) surveys, as attested by 69% of the articles (43/62); b) interviews according to 31% of respondents (19/62); and c) third-party observations according to 20% (12/60) of respondents. Moreover, in 33% of studies (20/62) no specific data collection method was reported. In addition, most ECB – the review shows - were evaluated with case studies according to 57% (35/62) or either pre-post or post-only with no comparison group, according to 14% (9/62) of the studies. Furthermore, the time period of data collection was either not reported, as confirmed by 27% of studies (17/62) or they ranged between “more than three years” and “one to two years according to 26% (17/62) and 16% (10/62) of the studies respectively.

---

<sup>77</sup> Encouraging reflective practices among organizations and governmental institutions being capacitated in evaluation is key to assessing the degree to which such ECD activities are contributing to change. Through the establishment of a cycle of action learning, ECD stakeholders will learn how to (i) identify a critical incident to reflect upon (experience); (ii) describe the experience (appraisal); (iii) examine the experience (analysis); (iv) interpret and draw lessons from the experience (discovery); (v) explore the alternatives and re-think future action (integration); and (vi) take action with new intent (Informed action) (Britton, 2010).

## CHAPTER 3

### METHODS

In this chapter, the methodological approach used (1) to explore the relationship between the terms ECB and ECD; (2) to identify the most critical national evaluation associations' organizational features and activities contributing to current ECD programming, and (3) to classify key ECD evaluative criteria and levels; is presented. The three questions addressed by this study were

1. To what extent is Evaluation Capacity Development (ECD) distinct from Evaluation Capacity Building (ECB) in international development contexts<sup>78</sup>?
  - a. What are the central attributes of ECD as compared to those of ECB?
  - b. What, if any, is the relationship between the terms ECB and ECD?
2. To what extent could ECD targeting in international development contexts become more inclusive in the future?
  - a. What is the current capacity of VOPE?
  - b. To what extent are VOPE currently involved in the conduct of evaluations and the promotion of an evaluative culture in their respective countries?

---

<sup>78</sup> It is noteworthy mentioning that the difference between Capacity Building and Capacity Development does not really exist in the French language. The most frequently used expression in French literatures is “*Renforcement des capacités*” and neither “building” (“*construction*” in French) or development (“*développement*” in French) is commonly used. For the purpose of this study, “*renforcement des capacités*” is equivalent to ECB. Therefore, the ECD model proposed by this study is applicable to and potentially beneficial to ECD programming in French-speaking contexts.

c. What are the factors characterizing the success or the failure of VOPE that should be taken into account in view of their involvement in ECD programs in international development contexts in the future?

3. How can, or how should, ECD best be evaluated?

a. What are the key criteria that need to be taken into account in order to assess ECD effectiveness?

b. At what level do the identified key ECD criteria need to be evaluated?

### Design

At a general level, the design of this study was both nonexperimental (subjects were not assigned to either a treatment or control group) and cross-sectional<sup>79</sup>. More specifically, the design of the study was a case study design (Chung, 2000; Morra Imas & Rist, 2009),<sup>80</sup> featuring both the simultaneous and sequential use of different qualitative methods<sup>81</sup> (Bamberger, 2000; Bamberger, Rugh and Mabry, 2012; Greene & Caracelli, 1997; Holton, Bates & Naquin, 2000; Stake, 1995; Van Rizin, 1995; Yin, 2002).

During the first phase, a systematic review of both ECB and ECD literature was conducted concurrently with a series of semi-structured interviews held with ECD specialists. During the second phase, a series of semi-structured interviews (informed by the findings of the first phase) were held with both ECD practitioners and VOPE members in three sub-Saharan countries (DRC, Niger, and South Africa). These interviews allowed the identification of new themes that were explored further during a second literature review phase. Based on both the information collected during the literature review phase as well as

---

<sup>79</sup> i.e., given the emerging nature of this study, those subjects included in the sample who were contacted more than once were asked to provide their feedback on a set of different issues at every follow up).

<sup>80</sup> Descriptive in nature, this study will be conducted in compliance with the Checklist for Reviewing the Quality of Case Study Report (Gilbert, 1982; GAO 1990; Yin, 1994) and the quality criteria to conducted Mixed-method evaluation.

<sup>81</sup> The use of concurrent qualitative methods appears to be well justified given that is emerging study, that is, a study based on an inductive approach.

the findings of in-country VOPE document reviews, two instruments (the SFAR Framework and the ECD Continuum Framework (Appendix F) were developed then submitted to a sample of ECD scholars, ECD practitioners, and VOPE members (both within and outside of Africa) for validation purposes. Through the involvement of several ECD practitioners and scholars, both the validity of findings and the accuracy of the qualitative information collected were enhanced.

Given the paucity of peer-reviewed literature not only on ECD but also on the evaluation of VOPE's role, a case study design was considered appropriate. Through a series of interviews and direct observations conducted during fieldwork, it was possible to gauge the type and quality of activities and processes inherent to three VOPEs in sub-Saharan Africa (the South African M&E Evaluation Association, the Niger M&E Evaluation Network, and the Congolese M&E Association), as described by their respective members. Likewise, through meetings with VOPE coordinators as well as other in-country ECD stakeholders, it became possible to both contextualize interviewees' responses and identify those organizational or environmental factors that appeared to either enable or hinder the VOPE involvement in the promotion of a in-country national evaluative culture<sup>82</sup>.

As the main focus of this study was on VOPE, no data collection was conducted among CSOs conducting evaluations. Likewise, only a few semi-structured interviews were conducted among local evaluation service providers (see Box 1).

---

<sup>82</sup> This is all the more relevant at a time when VOPE are increasingly formalizing their structures across Africa and the need is there for the sharing of lessons learned among VOPE to this date, this study will allow for "understand[ing] what meaning people give" (Bamberger, Rugh & Mabry, 2012, pp.597) to ECD programming and "how they perceive its purpose, their attitudes and expectations, how they respond to it, how their response is affected by contextual factors, and what effect it has on them" (Bamberger, Rugh & Mabry, 2012, p. 598).



### **Box. 1 Why Focusing only on VOPE?**

The main focus of this study was on VOPE and no specific data collection was conducted among any of the NGO conducting evaluation work in the three countries where the case studies were developed. The only non-VOPE experience presented in this study was that of three different evaluation service providers (which VOPE might want to establish closer links with in the future): one government training agency in South Africa (PALAMA), one private institute in Niger (L'Institut de Stratégie, d'Evaluation et de Prospectives) and one university department in Niger currently developing the curriculum for a Master's degree in M&E (Department of Sociology at the Social Science School). Despite the seemingly reductive nature of the study's scope, the justification for focusing exclusively on VOPE was double-fold.

First, unlike NGO that had been receiving M&E training or financial support to deliver M&E capacity building programs to a variety of local stakeholders over the last decade, VOPE (mostly due to their informal status) had never been able to benefit from such resources in the past. Therefore, at a time when a large number of VOPE were trying to acquire a formal status and a new set of ECD activities were being planned to enhance the status of evaluation knowledge and practice in many countries, it appeared quite relevant to explore what complementary and innovative role VOPE could play as part of the ongoing ECD endeavors. What was specifically important was to understand what VOPE needs and capacity were and what needed to be done in order to turn them from passive receivers of capacity building activities to "active" ECD agents (e.g., providers of evaluation services).

Second, unlike NGO that have often been viewed by governments as untrustworthy "anti-establishment" organizations, VOPE, characterized in this study as "sphere-crossing" entities with a very eclectic membership, seemed more adequate to foster a more constructive dialogue between the government and a variety of stakeholders outside of the governmental sphere. Far from attacking governmental institutions, VOPE (which in most cases include government officials among their members) seemed to be the ideal promoters of a more inclusive and long-term ECD planning.

Due to the emergent nature of the study,<sup>83</sup> constant-comparative methods were a central element of the study's design and execution (Strauss & Corbin, 1990). Through a dialogic process, the plan, data collection and analysis were continuously enhanced

---

<sup>83</sup> An emergent study is justified when (i) prior knowledge of a given phenomenon is missing, (ii) methodological tools are inappropriate; and (iii) situational control is inadequate for applying a pre-ordinate design.

throughout the conduct of the study. Similarly, preliminary interpretations were used to sharpen the focus for subsequent data collection phases.

In order to address the three main research questions, three main methods were used:

1. A systematic literature review
2. Semi-structured interviews
3. VOPE documentation review

During the last phase of data collection, a series of online follow-up interviews and offline discussions with VOPE and ECD specialists were held.

#### Data Collection Methods

A systemic literature review of (1) articles published in peer-reviewed journals, (2) specialized books, and (3) academic-level work published in grey literature available online; was conducted.

#### Sample

A non-probability purposive sample was used to locate the literature currently available on the three key study topics: (1) the relationship between ECB and ECD; (2) the VOPE effective contributions to current ECD programming; and (3) ECD evaluative frameworks.

Consistent with the emerging nature of the study, the literature review was conducted in two phases. During a first phase, a sample of articles and reports ( $N=194$ ) was identified based on an electronic search for seven key terms (e.g., ECB, ECD, ECB

evaluation, ECD evaluation, National Evaluation Associations, VOPE; and ECD evaluation). More specifically, the online search was conducted on the following:

- a) Three peer-reviewed evaluation journal indexes (the *American Journal of Evaluation*, *New Directions for Evaluation* and the *Canadian Journal of Program Evaluation*);
- b) One database containing a large variety of social science journals (the Social Science Citation Index);
- c) Websites of over a dozen leading international development agencies and foundations involved in ECB and ECD programming (e.g., World Bank ECD Working Papers, OECD, UN evaluation offices websites, Rockefeller Foundation; European donors' websites – e.g., Department for International Development (Dfid), the Norwegian Agency for Development Cooperation (NORAD), the Swedish International Development Agency (SIDA); the USAID Office of Learning Evaluation and Research (PPL/LER)<sup>84</sup>;
- d) Five websites of organizations and communities of practice (CoP) specialized in ECB and ECD (Evaluation Capacity Development Group, European Centre for Development Policy Management; EVALTALK, XCeval, and Pelican Initiative).

## Procedure

In order to identify which articles and reports were to be included in the systematic literature review, a checklist was used to determine whether the articles and reports, identified through the electronic search, met at least one of the following themes:

- a) They provided definitions of ECB and/or ECD

---

<sup>84</sup> The list of all the websites consulted can be found in one of the Annexes added to the final copy of the dissertation.

- b) They specifically described VOPE and/or national evaluation associations and internal processes and activities;
- c) They mentioned specific variables or methods used to evaluate ECB and ECD programs

Only those articles and reports that touched upon at least one of the topics delineated above ( $n = 104$ ) were included in the systematic review. This number of articles included during the first phase of the systematic literature review was complemented by an additional number of a peer-reviewed articles and grey literature – especially indigenous literature not available online ( $n=93$ ) that were identified based on (1) the references included in the sample of literature consulted during the first review phase; and (2) the recommendations made by the ECD scholars and in-country practitioners interviewed during the data collection phase (snowball sampling).

The search was exhausted once the literature identified during the second phase of the review (and more specifically focusing on CB and CD) started yielding definitions of terms as well as processes evaluative frameworks that overlapped with those already identified during the earlier stage of the review.

#### Data Processing and Analysis

The analysis of the literature review consisted of three phases. First, once a sample of cases satisfying one or all the inclusion criteria was identified, the articles and reports included in it were divided in three different groups based on the specific research question which they were the most closely linked to<sup>85</sup>. During this first phase, the unique

---

<sup>85</sup> As some articles and reports were relevant to all three key research questions, the inclusion in one group was not exclusive. As a result, several cases were included in all three groups.

contributions provided by each of the articles or reports included in the sample were commented on and placed in the narrative found in Chapter IV.

Second, a comparison of the qualitative research findings was conducted across cases using a grounded theory approach and developing a multi-coding system to group them (Miles and Huberman 1994; Yin, 2002).

Third, findings across cases were compared in order to identify the outcomes and pathway of ECD.

In order to address sub-question 1.2 on the relationship existing between ECB and ECD, the characteristics associated with each of them, as described in the literature available were summarized and collapsed into a table as shown in Table 6. Given the emerging nature of this study, the original list of categories developed to compare ECB and ECD was revised in the course of study.

The list of ECB and ECD evaluative criteria included in the original framework presented above was revised and expanded based on the second literature review phase as well as the feedback provided by a variety of ECD practitioners at three different times, thus enhancing the face validity of the instrument (Appendix F).

Table 6

*ECB/ECD Matrix*

Type of Programming Characteristic	ECB	ECD
Goal		
Objective		
Key Assumptions		
Ideal Strategy		
Type of Accountability		
Scope of activities		
Key groups to capacitate		
Venue of capacitation		
Level of difficulty to evaluate		
Type of Evaluation Champions		
Degree of personalization of the evaluation function		
Status of incentives availability		
Degree of Ecological Awareness		
Timeframe		
Level of implementation costs		
Type of funding source		
Degree of financial transparency		
Modality of Capacity providers selection		
Nationality of capacity providers		
Level of Ownership		
Level of Sustainability		
Innovation Level		
Level of ability to address equity and vulnerability		

#### Semi-structured Interviews

Semi-structured interviews were conducted both in person and online to frame this study within the broader ECB and ECD discourse. More specifically, semi-structured interviews were conducted with four different groups:

- a. ECD leading scholars;

- b. ECD specialists working within international development agencies (multilateral, bilateral and foundations);
- c. Representatives of VOPE other than the ones visited during the field work in Africa
- d. VOPE coordinators and other members of VOPE visited during the field data collection in Africa.

One common methodology was used to develop and administer the semi-structured interviews held with the first three groups (also referred to in this study as Advisory Panels).

#### Methodology for Semi-structured interviews with the Advisor Panels

The objective of the semi-structured interviews (Appendix B) was to analyze further and validate some of the study’s emerging themes and preliminary conclusions at multiple points in time during the study. More specifically, the preliminary findings and conclusions of this study were validated through a multi-stage participatory process (see Figure 6).

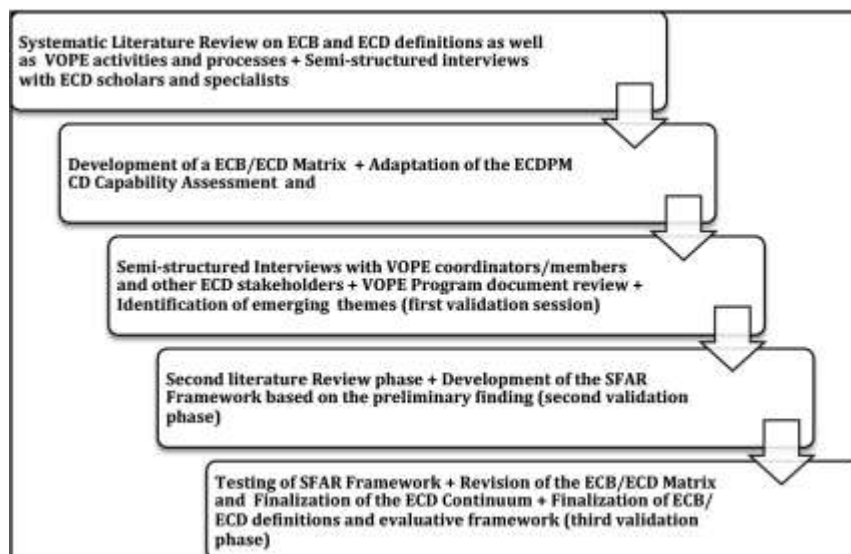


Figure 6. Sequencing of Qualitative Methods

## Sample

A non-probability purposive sample was used sample from the population of ECD scholars and ECD specialists working in international organizations, national development agencies and foundations supporting ECB and ECD programming.

Professors, researchers, and practitioners with an in-depth knowledge of ECD and its concrete application in the field and whose names had either been found in some of the consulted literature articles as well as recommended by peers, were included in the group of ECD scholars ( $n = 10$ )

Seniors staff with extensive experience in ECD programming working with a variety of organizations (e.g., OECD, the World Bank, the Asian Development Bank, African Development Bank, the United Nations Development Programme, UNICEF, the Rockefeller Foundation and UNWomen) were included in the group of representatives from the donor community ( $n = 15$ ). Furthermore, ECD practitioners that worked outside of the three countries where the case studies were developed, were interviewed ( $n = 9$ ).

Overall, a total of 78 people were interviewed in the course of field data collection in Niger, DRC and South Africa (see Table 8).

Semi-structured interviews were conducted with 27 ECD practitioners in DRC (3 from the ACoSE Coordination Team, 3 from SENAREC, a national training institution, 2 from UNFPA, 10 from the government, 9 from Civil Society).

In South Africa, interviews were conducted with 16 ECD practitioners (2 from DPME, 1 from the Public Service Commission, 2 from the Auditor's General Office, 3 from the SAMEA Coordinating Committee, 1 from a National training institution, 7 from NGOs).



In Niger, interviews were conducted with 35 ECD practitioners (2 from the ReNSE Coordination Committee, 2 from the Ministry of Planning, 11 from other Ministries, 14 from CSO and implementing agencies, 7 from international organizations)

An additional number of evaluation practitioners, equally distributed across Africa, Latin America, the US and Europe ( $n = 9$ ), were also interviewed during the final phase of data collection.

Table 7

*Respondents' Profile*

<b>AREA</b>	<b>Sector</b>	<b>Frequency</b>
GLOBAL	ECD Scholars	10
	ECD Officers in Multilateral organizations	14
	ECD Officers in Foundations	1
	ECD practitioners	9
	Sub-total	34
DRC	VOPE Coordination Team	3
	National Training Institutions	3
	Donors	2
	Government	10
	Civil Society	9
	South Africa	16
	Sub-total	27
Niger	VOPE Coordination Team	2
	National Training Institutions	1
	International Organizations	7
	Government	13
	Civil Society	14
	Sub-total	37
South Africa	Government	5
	VOPE Coordinating Team	3
	National Training Institution	1
	Civil Society	7
	Sub total	16
	<b>Total</b>	<b>114</b>

The group of representatives of VOPE other than those visited during fieldwork included coordinators of VOPE particularly involved in the promotion of evaluation both at the national and regional level in Africa, Latin America and South-east Asia ( $n = 52$ ).

#### Instrumentation

The list of questions, used during the initial round of phone- or in person semi-structured interviews with the members of the Advisory Panels, included:

1. How would you characterize Evaluation Capacity Building and Evaluation Capacity Development?
2. Is there difference between the two terms? If so, what is the difference?
3. What is VOPE role in the conduct of evaluation and/or the promotion of an evaluation culture?
4. To what extent could VOPE get more involved in national ECD planning processes in the future?
5. Would you be able to recommend other scholars that work on ECD?

Follow-up interviews, either over the phone or in person as well as offline discussions (depending on respondents' availability), were held throughout the study to touch upon some of the emerging themes identified at the end of the first data collection phase (e.g., the relevance of long-term champions, the impact of missing incentives at all levels, the inevitable VOPE sustainability challenges following their formalization process) as well as to discuss some of the preliminary conclusions (e.g., as those included in the Draft SFAR Framework or ECD Continuum reached at the end of the second data collection phase.

## Procedure

The study adhered to all protocols established by the Human Subjects Institutional Review Boards (HSIRB) (see Appendix C). The individuals invited to take part in the study were contacted through an introductory e-mail explaining both the purpose of the study and the type of involvement that their participation in the research would entail (that is, the participation in off-line discussions three times during the following 4 months).

Dear ECD Practitioners/Scholar:

I hope this note finds you well. My name is Michele Tarsilla and I am a Ph.D. Candidate in Interdisciplinary Evaluation at Western Michigan University. I have been working on international development evaluation for over 10 years and I am currently conducting a global research on Evaluation Capacity Development (ECD) in international contexts (this is my doctoral dissertation at Western Michigan University and both Michael Bamberger and Jim Rugh are on my committee). I will be in Africa over the next three months and I am planning to develop two case studies on ECD: one in South Africa and one in the Democratic Republic of Congo.

As part of this research, I am setting up an advisory panel, including several world's leading experts on EDC. As I understand that you and your agency have a special interest in this topic and in light of your wide experience in this area, I would like to invite you to be part of this panel. The idea is for me to lead a non-simultaneous online discussion among ECD gurus in the international donor community once or twice a month over the next 4 months. The online discussion would be informed by my preliminary findings in the field. Please let me know if this would be of interest to you.

Understanding what challenges (if any) your agency is currently facing in the ECD area and gaining a better sense of the existing knowledge gaps that my research could hopefully contribute to shed some light on would be very important. I am well aware that you are extremely busy. However, I am confident that getting you on board could definitely help shape the data collection instruments that I am planning to use in the field. It would also allow me better understand how my work could contribute to your agency's current ECD endeavors.

I would be honored to get a chance to speak with you soon. If you are interested in participating in the study, please reply favorably to my personal e-mail address: [michele.tarsilla@wmich.edu](mailto:michele.tarsilla@wmich.edu) I look forward to hearing back from you.

If you decide to participate, you will receive a synthesis of the findings of the study.

Warm regards,

Michele (Italian for Michael)

Michele Tarsilla

PhD. Candidate in Interdisciplinary Evaluation at Western Michigan University

Phone: +1 (202) XXX-XXXX<sup>86</sup>

### Analysis

For all the responses to the open-ended questions included in the semi-structured interview and the offline discussions as well as for all the themes identified in the course of the literature review, the preliminary findings were finalized inductively. Content analysis was conducted (thematic analysis) and a special effort was made to situate the recurrent patterns among responses within the broader context where such feedback/answers were developed (discourse analysis; (Miles & Huberman, 1994). Put simply, qualitative data was interpreted for the purpose of explanation.

Through the use of thematic analysis, common patterns in respondents' feedback were identified concurrently and, through an adequate coding exercise, these were translated in quantitative terms.

### Use of Triangulation and Scenario of Convergence/Divergence

After a careful scanning of results and the identification of main findings and patterns, the extent to which different data sources point toward the same need (in which

---

<sup>86</sup> The number originally included in the e-mail sent out to potential interviewees was omitted in the final copy of this dissertation for the sake of confidentiality.

case, the results will be presented as complimentary) or not (Moseley & Heaney, 1994) was assessed.

In case of disagreement among sources, the advisory panel was asked to contribute their feedback and to assist with disagreements (Datta, 1997). The advisory panel was also instrumental in certifying the quality and integrity of the research (member checking), besides serving as communicators/advocates of the study in international settings and with their respective professional networks (fourth general evaluation). In any case, the strongest data source in regard to clarity of outcome and strength of implementation will be identified.

#### Semi-structured Interviews with VOPE Coordinators and Other VOPE Members

The semi-structured interviews with VOPE coordinators and other VOPE members aimed at assessing VOPE characteristics (e.g., organizational development and technical capacity) as well as their role in conducting evaluation and promoting a national evaluative culture were conducted based on a slight variation of the methodology used for the semi-structured interviews with the other three groups of interviewees.

#### Sample

A three-stage sampling was used to identify the national evaluation associations' members to interview as part of this study.

##### *1) First Stage*

Data were collected in three purposefully selected countries. Two of the three countries (South Africa and the Democratic Republic of Congo) were identified based on extreme or outlier case sampling (Flyvbjerg, 2006):

- a) South Africa had one of the most vibrant national evaluation association in sub-Saharan Africa (SAMEA was established in 2004) and had created an ad hoc M&E Department under the aegis of the President;
- b) The Democratic Republic of Congo (DRC) had an evaluation association that was still struggling to gain a formalized status and had a limited number of channels available through which government accountability could be enhanced.

Niger (the third case study) was identified based on a critical case sampling:

- c) Recipient of a large financial support for the development of evaluation communities of practice evaluation and management for development results initiatives since 1999, the VOPE was confronted with some critical sustainability issues at the onset of this study. As VOPE sustainability is a common area of concern for ECD programs, an in-depth review of the national VOPE was recommended by several experts - consulted at the onset of the study—on the grounds that the Niger case would be informative to both enhance ECD performance over time in the absence of continued funding and incorporate some strategies enhancing ECD into ECD design. Put simply, analyzing the current challenges faced by a country that had been regarded as a particularly ECD success story in the past but was no longer in a good standing with development partners today was believed to be useful to all other VOPE who were likely to a) experience an increase in ECD funding in the near future and b) face the same challenges

experienced by the VOPE in Niger. The assumption here is that “if it happens there, it will happen everywhere.”

*2) Second stage*

Within each country, a plurality of stakeholders involved in evaluation (identified through both purposeful and snowball sampling) in the three countries’ capitals (Pretoria/Johannesburg in South Africa; Kinshasa in DRC; and Niamey in Niger) were identified at three different levels (Heider, 2011) between March and June 2012

- a) Level 1: VOPE Coordinators and other VOPE members
- b) Level 2: Civil society representatives; academia and other entities involved in ECD programs outside of government; (this would include development partners;
- c) Level 3: government officials involved in ECD programs

*3) Third stage*

Within each level, a purposive sample of subjects was identified.

Cognizant of the extractive nature of past research conducted in development contexts, data collection was anticipated by a presentation delivered by the main author of this study in all three countries. The presentation topic was selected collectively by VOPE members before the start of the data collection phase. In the case of Niger, it was a presentation of impact evaluation methods in the public sector.

## Instrumentation

Semi-structured interviews were conducted with VOPE coordinators and a smaller purposeful sample of VOPE members in the DRC, Niger and South Africa to assess VOPE capacity and processes. The interview protocol drew on the Core Capability Assessment tool developed by the European Center For Development Policy Management (ECDPM) in 2011 and validated based on the conduct of 16 global CD case studies. The original questions included in this questionnaire were grouped in 5 different categories (Appendix D):

- a) Capacity to Commit and Act:
- b) Capacity to Generate Development Results:
- c) Capacity to Relate:
- d) Capacity to React:
- e) Capacity to Integrate:

Building on this existing instrument and adapting it to VOPE, a new version of the tool was developed and called VOPE Capacity Assessment Tool (see Appendix E). The tool allowed a better understanding of critical VOPE organizational features that need to be taken into account before implementing activities more specifically aimed at enhancing VOPE evaluation capacity. ECD and ECB are tightly connected to organizational development. The higher the organizational capacity, the higher the organization's propensity to learn (organizational readiness for evaluation).



## Procedure

The items included in the VOPE Capacity Assessment were asked in the course of the semi-structured interview conducted with the VOPE coordinator and other VOPE members. The key preliminary findings were discussed with the respondents twice throughout the study.

## Analysis

Based on the responses provided by their coordinators and members, the three VOPE of interest were compared across a series of evaluative criteria (see Table 8) that had been demonstrated to be particularly relevant both in the literature review, the semi-structured interviews conducted with ECD Scholars as well as a preliminary content analysis of the responses to the VOPE Capacity Assessment (conducted during the first phase).

Table 8

### *VOPE Organizational Processes: Comparative Checklist*

Country	DRC	Niger	South Africa
Criteria			
Historical Development			
Organizational Development Phase			
Membership Diversity			
Compliance with internal government rules			
Degree of Internal Networking			
Resilience			
Leadership Type			
Ownership			

- Development Phase (Pioneer, Differentiated, Integrated), (Ubels et al., 2010, p.123);
- Resilience (ability to weather period of crisis- internally or externally triggered),
- Leadership (concentrated, decentralized and network/shared,
- Sovereignty (self-reliance and local ownership),

Similarly, In order to assess the capacity of the different national evaluation associations presented in the three country case studies discussed in the this chapter, the two following instrument will be used: the adapted version of the Five Core Capabilities (Fowler & Ubels, 2010). In order to summarize some of the key findings, the table below (see Table 9) was used.

For all answers to open-ended questions included in the offline discussion as well as for all responses to the semi-structured key informant interviews and all feedback provided in the course, preliminary findings were finalized inductively. Content Analysis was conducted (thematic analysis) and a special effort was made to situate recurrent patterns

Table 9

*VOPE Main Activity Main Objectives*

Country	DRC	Niger	South Africa
Criteria			
Education/training			
Research			
Policy Advisory Work			
Advocacy			
Promote evaluation culture			
Defend the interest of evaluators as professionals networks			
Networks			

\* Low; \*\* Medium; \*\*\* High

among responses within the broader context where such feedback/answers were developed (Discourse Analysis). (Miles & Huberman, 1984). Put simply, the qualitative data or tacit

knowledge (Polanyi, 1958) was interpreted for the purpose of explanation (von Wring, 1987).

Through the use of thematic analysis, common patterns in respondents' feedback were identified and, through an adequate coding exercise, these were translated in quantitative terms (Schofield, 2002). They also served as the basis for follow-up interviews.

Not everybody in the sample ( $N=114$ ) was asked to answer the same number of questions. ECD scholars and specialists working in the headquarters of bilateral and multilateral organizations (Bangkok, New York, Tunis, Washington DC) ( $n=24$ ) were asked a more comprehensive set of questions (including some more conceptual ones) than the in-country ECD practitioners (e.g., VOPE coordinators and representatives from academia and the private sector) ( $n=90$ ). Therefore, in order to prevent any ambiguity in reporting the findings of this study, for each of the percentages presented, the corresponding sample was clearly indicated. For example, in order to indicate the 72% of the in-country practitioners (and not of the whole sample) mentioned pre- and post-test as the most frequent methods to evaluate trainings' effectiveness, the corresponding statistics reported in this study was 72% ( $64/90$ ), where 64 was the actual number of respondents and 90 the sample group of in-country practitioners.

## Review of VOPE Documentation

### Sample

Consistent with the VOPE Capacity Assessment's diagnostic purpose, a purposive sample of program documents was reviewed among all those provided by each VOPE in DRC, Niger and South Africa in the course of the field data collection.

## Instrumentation

A standardized list of program documents from which the sample was selected was used before arriving in the country. This included the VOPE Action Plans, Job Descriptions, Budgets, Proposals, Evaluation Reports, National Evaluation Policy, Annual Reports, Activity Reports, Conference presentations. During the review of the documents provided by each VOPE coordinator, the critical documents that were missing from the list were identified and a follow up with the VOPE coordinator to receive them was made.

## Procedure

The list was submitted to the VOPE Coordinator and other VOPE members following the first semi-structured interview. Several e-mail follow-ups were made to solicit the provision of the requested materials within a month from the interview.

## Analysis

A content analysis all the documents provided was conducted for two main reasons. On the one hand, it allowed triangulating some of the emerging findings from the literature review and the semi-structured interview. On the other hand, it allowed identifying new themes that it was worth exploring further through the online survey or the follow-up exchange with the ECD scholars and practitioners.

## CHAPTER 4

### FINDINGS FROM THE SYSTEMATIC LITERATURE REVIEW

In this chapter, the findings are presented from interviews held with a variety of ECB and ECD practitioners working in international development contexts, as well as from a systematic review of literature on ECB, ECD, and other closely related terms.

#### The Relationship between ECB and ECD

##### Findings from the Systematic Literature Review

In an effort to better understand the relationship between ECB and ECD (see research question 1), the most popular definitions of both terms available in peer-reviewed and grey literature were identified and commented on. Likewise, the definitions of other terms closely related to both ECB and ECD (e.g. capacity, capability, competency, and capacity development) were collected, analyzed, and presented in this chapter.

Through this systematic review of key ECB- and ECD-related terms, two primary objectives were fulfilled. First, the most critical developments that occurred in both the conceptualization and interpretation of the terms ECB and ECD over the last decade were identified. Second, through the identification of ECB and ECD central attributes, a list of core elements of successful ECD programs that one might want to take into account when designing and evaluating ECD activities in the future was developed.

The terms presented in this section are presented in the following order: capacity, capability, competencies, capacity development, evaluation capacity building, and evaluation capacity development. For these terms, definitions are provided in chronological order.

Three summaries of the numerous definitions are provided for three of the six terms:

Capacity (Appendix G), Capability (Appendix H), and Capacity Development (Appendix I).

## Capacity

Twelve main definitions of capacity were identified in international development literature and adequately supported by a variety of stakeholders in the field.

Capacity (1): “The ability to carry out stated objectives” (LaFond & Brown, 2003, quoting Goodman, 1998, p. 7).

In this first definition, capacity was described as a function of a rather linear and result-oriented planning of actions (instrumental or functional perspective of capacity). This definition, which was often ascribed to the Results-Based Management (RBM) logic a decade ago, was particularly popular among 70% (33/47) of the planners and managers interviewed in the field. Interestingly, among the 88% (22/25) of planners and managers agreeing with this definition, the majority resided in Niger.

Capacity (2): “The ability of individuals, institutions, and societies to perform functions, solve problems, and set and achieve objectives in a sustainable manner” (UNDP, 1998).

Although this second definition was built upon the functional perspective of capacity associated with the prior definition (capacity was described as a means through which a certain number of functions could be performed), it was innovative for two different reasons. First, a definition of to whom the capacity belonged (individuals, institutions and

societies) was provided, and the foundations were laid for the three-level ECD analytical framework (individuals, organizations, and enabling environment) that would be widely used and revised by numerous ECB and ECD scholars and practitioners during the following decade (see Chapter 2). Second, through this definition two important principles were introduced: the need for sustainable results, which 74% (85/114) of stakeholders identified as a necessary attribute of successful capacity strengthening programs, and the need for being responsive to emerging issues, mentioned by 31% (36/114) of stakeholders.

Capacity (3): “An organization’s potential to perform—its ability to successfully apply its skills and resources to accomplish its goals and satisfy its stakeholders’ expectations” (Lusthaus et al., 2003, p. 3).

In this third definition, which was focused on only one of three levels (organizational levels) that had been referred to earlier, three new attributes of capacity were identified: a) the latent nature of capacity; that is, its availability to pursue a continually evolving set of objectives (as mentioned by an M&E officer working for one UN agency in the DRC: “Capacity is there regardless of whether or not you have development programs targeting it”); b) the technical nature of capacity (as attested by one government official in Niger: “The application of concrete skills is a sign of capacity”); and c) the accountability purpose and participatory nature inherent to the concept of capacity, as confirmed by a government official interviewed in Pretoria:

Enhanced capacity is a means to address the expectations of an organization’s stakeholders. That is true at the broader national level, as well. For instance, it was because of this awareness of the link between capacity and stakeholders’ expectation that our department pushed for the development of initiatives, such as the Citizens’ Hotline, that aimed at responding directly to citizens’ needs for better public services.

Capacity (4): “[The] potential to perform” (Horton, Alexaki, Bennett-Lartey, Brice and Campilan et al., 2003, p. 7). In this concise definition the latent nature of capacity presented in the prior definition was reiterated.

Capacity (5): “The potential for using resources effectively and maintaining gains in performance with gradually reduced levels of external support” (LaFond & Brown, 2003 p.13). In this fifth definition that was built upon the prior two, two of the most widely discussed themes in international development today were introduced: organizational/programmatic sustainability. and value for money. Both of these themes were mentioned as an important objective of capacity strengthening programs, respectively by 76% (61/80) and 49% (56/114) of the stakeholders met in the field.

Capacity (6): “The ability of people, organizations and society as a whole to manage their affairs successfully” (OECD, 2006a, p. 9).

Although this definition seemed to be aligned with the functional view of capacity provided in some of the earlier definitions, it was innovative in that it provided a more holistic view of capacity. Thus, it placed a renewed emphasis over the collective nature of capacity (e.g., the more encompassing term “people” replaced “individuals”; the more unifying term “society” replaced “societies”).

Capacity (7):

[It] refers to people acting together to take control over their own lives in some fashion. Capacity is a potential state. It is elusive and transient. It is about latent as opposed to kinetic energy. Performance, in contrast, is about execution and implementation or the result of the application/use of capacity. Given this latent quality, capacity is dependent to a large degree on intangibles. It is thus hard to induce, manage and measure. As a state or condition it can disappear quickly, particularly in smaller, more vulnerable structures. This potential state may require the use of different approaches to its development, management, assessment and monitoring. (Morgan, 2006, p. 6)



This definition set itself apart from the previous ones for several reasons. First, in going beyond the instrumental view of capacity presented in the first definition, capacity was legitimized as an existing feature of both individuals and collective entities, regardless of their actual use<sup>87</sup> towards the attainment of pre-states objectives (as argued by the third definition). Second, a more dynamic perspective of capacity was introduced that specifically pointed to the uncertainties and unpredictability of the conditions which capacity depended on. Third, it was suggested in the definition that capacity could be developed, managed, and assessed but not built, as confirmed by 63% (49/78) of the stakeholders interviewed in the field.

Capacity (8):

[...] A potential state of performance, the emergent combination of individual competencies, collective capabilities, assets and relationships that enables a human system to create a value (Morgan & Baser, 2008, p. 38)

The emerging nature of capacity was reiterated by this spontaneous and naturalistic definition. In addition, other terms distinct from capacity and yet closely related to it (competencies and capabilities) were introduced for the first time. Furthermore, a systemic (capacity is relational) and value-based perspective on capacity was introduced by this definition for the first time.

Capacity (9):

A standing attribute of any organization, it could either consist of a) basic functionalities or b) organizational talents. On the one hand, basic functionalities represent the 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. Basic functionalities are the minimal systems,

---

<sup>87</sup> The use of finding does not always happen immediately after the release of the evaluation report and does not always translate into a specific action. Otherwise said, utility might emerge at the macro level and not be time specific (Henry & Mark, 2003).

infrastructure, resources, collective ability and coherence needed for an organization to perform consistently well over time. On the other hand, organizational talents represent what one does, thinks or feels consistently (well) and can put to productive use. As such, they are distinct from individual knowledge (what one knows) and skills (what one knows how to do). Organizational talents represent 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. Organizational talents are about highly effective, creative, innovative organizations whose synergies are well beyond the sum of their individual parts. [...] 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. (Ortiz & Taylor, 2009, p. 32)

Through this definition, the tension, characterizing all the prior definitions, between the latent nature of capacity and its effective articulation and use appeared to be finally solved. Through the introduction of the concept of standing capacity, the need was identified for systems to develop permanent and effective qualities and resources, enabling them to react to both predictable and unpredictable situations<sup>88</sup>. A corollary of such definition was that any comprehensive initiative aimed at enhancing countries' evaluation capacity would need to assist with the development of organizational talents and not simply the application of standing capacity (e.g., skills usable right after training delivery), as it was the focus instead of shorter-term initiatives.

Capacity (10): "The emergent combination of attributes, assets, capabilities and relationship that enables a human system to perform, survive and self-renew" (Fowler & Ubels, 2010, p. 34).

---

<sup>88</sup> The need for identifying prior development objectives and verifying the compliance of the project with envisaged results as the basis for estimating its value is increasingly common in evaluation: "Despite the high levels of uncertainty within the development environment, there is a pressure to be able to predict and to appear infallible and pursue the so-called matrix culture" (Pasteur, 2006: 35).

Although the emerging nature of capacity as well as the relevance of assets, capabilities, and competencies presented earlier was confirmed by this definition, a new important concept was introduced by it: capacity as an internal and critical feature of a human system. According to this definition, capacity was not simply what would facilitate the attainment of any stated objective but also, and foremost, the basis for survival and self-renewal, as confirmed by seventy-two percent (72%) (58/80) of the stakeholders interviewed in the field. As stated by the director of a private evaluation training institution in Niger: “Capacity is about development. If capacity stagnates, countries in Africa will never solve their problems and will keep depending on donors’ support.”

Capacity (11): “The ability to define and realize goals, where defining goals entails identifying and understanding problems, analyzing the situation, and formulating possible strategies and actions for response” (UNICEF, 2010, p. 24).

Through this definition, some breadth was certainly added to the operationalization of the concept of capacity. However, despite the close affiliation of such definition with the programmatic perspective of capacity presented earlier (goal setting, situational analysis, planning of strategies and actions in response to identified issues), not too much information was provided in it on who were the actors involved in this systematic process and at what level they were operating.

Capacity (12):

Attribute of people, individual organizations and groups of organizations. Capacity is shaped by, adapts to, and reacts to external factors and actors. It includes skills, systems, processes, ability to relate to others (internally and externally), leadership, values, formal and informal norms, as well as loyalties, ambitions and power. Thus, capacity development is a change process modifying some of these factors, or their configuration. (Boesen, 2010, p. 147)

In this last definition, more comprehensive than the earlier ones, the collective dimension of capacity was reiterated and a number of new elements were introduced. First, through the reference to groups of organizations (rather than society) as the highest level of capacity applicability, the discourse on capacity was brought to a lower and more manageable level of analysis. Second, through the emphasis on the influence that internal and external factors had on capacity, the relevance of context was acknowledged in the definition and a more thorough analysis of internal dynamic and external processes was called for. Third, through the use of such concepts as leadership, loyalty, ambition, and power, the relational, psychological, and social attributes of capacity were reiterated.

#### Additional Remarks on Capacity

Some further specifications of capacity exist that are somewhat complementary to the broader definitions presented above.

Capacity, for instance, could be generic or specific (OECD, 2006). *Generic capacities* include the ability to plan and manage organizational changes and service improvements; *specific capacities* are those developed and practiced in critical fields, such as public financial management or trade negotiation.

Likewise, according to the definition of capacity developed by several IDRC evaluation officers (Earl, Carden & Smutylo; 2001), an organization's capacities could be distinguished between *operation capacities* (the capacities needed by an organization to carry out as part of its day-to-day activities) and *adaptive capacities* (the capacities needed by an organization to learn and change in response to changing circumstances). This distinction seemed quite relevant for the sake of this study as it reiterated the argument that capacity was functional to learning, one of the primary purposes most frequently associated with the

evaluation function, as predicated by a large number of evaluation policies in use by bilateral and multilateral aid organizations.

Furthermore, capacities could be divided between hard (e.g., infrastructure, technology, and finances) and soft (e.g., “human and organizational capacities, or social capital of the organization, [including] management information systems, and procedures for planning and evaluation”) (Horton et al., 2003 p. 34). In addition, soft capacities could be distinguished even further between “tangible” (the ones mentioned above), and “intangible,” that is, the ability of an organization to function as a resilient, strategic and autonomous entity (Kaplan, 1999) as well as to “commit and engage, adapt and self renew, relate and attract, and balance diversity and coherence” (Baser & Morgan, 2008; Morgan, 2006). These last capacities, in addition to the one “to carry out technical service delivery and logistical tasks,” were particularly important within the scope of this study as they were integral components of the Five Capabilities model developed by the European Centre for Development Policy Management (Engel et al., 2007; Horton et al., 2003) whose adapted version was used during the field data collection to measure the capability of the three VOPE in the DRC, Niger, and South Africa. Overall, the concept of capacity as “improved performance” (typical of the earlier definitions that would be later associated with a reductionist thinking perspective) was increasingly replaced by a multi-dimensional definition of capacity reiterated by the last three definitions<sup>89</sup>.

---

<sup>89</sup> From a theoretical perspective, such definition is well aligned with the “complex adaptive systems (CAS) thinking” (Watson, 2010) and the concept of Presencing (Magruder-Watkins and Mohr, 2001). On a more pragmatic level, such definition has been openly embraced by both the Australian Agency for International Development (AusAID) and the German Federal Ministry for Economic Development Cooperation (BMZ), which reported to have abandoned logical frameworks in planning programs. As a result, input, activities and outputs are worked out by the implementation teams and their partners and logframes have been replaced by results chains (sketches of how change is envisaged). Similarly, UNDP concluded that RBM was not used effectively (UNDP 2007) and Dfid came up with a multi-year outcome based Challenge grants, a more flexible and responsive funding option.

## Capability

Often confounded with the concept of capacity, capability became a relevant and increasingly recurrent element used within the broader ECB and ECD discourse as of the mid 1990s. The distinctive features of the term were identified by the six definitions presented below.

### Capability (1):

An institutional feature including a) a variety of services to be performed, b) the demand for these services, c) capabilities (skills in necessary quantity and quality) to meet demand, d) hardware (including facilities, premises and instruments), e) methodology, f) legislation, and g) management and coordination functions. (UNIDO, 1990)

Although the holistic definitions of capacity presented earlier seemed to be echoed by this comprehensive definition of capability, three main differences between the meaning of the term and any prior attempt to define capacity were identified. First, through the emphasis on the unique applicability of the concept to institutions (not individuals or the society more in general), the collective and tangible nature of the term was highlighted. Second, given both the use of the concept of supply and demand for services that had dominated the ECB discourses in the 1990s and 2000s and the appreciation for both the management and coordination function associated with the term, capability was defined as a relational feature of any given organization. Third, due to the recognition the importance of skills and facilities (micro-level variables) and legislation (macro-level variable), the need for an adequate assessment of the capability level within any given institution through an environmental scan seemed to be suggested.

Capability (2): “The individuals’ ability to do something” (Sen, 1999).

This second definition was quite distinct from all the others provided in this section as the individual (rather than collective) nature of capability was highlighted, without being tied to any specific objective, individual, or contextual variable or function.

Capabilities (3):

The collective ability of a group or a system to do something either inside or outside the system. The collective skills involved may be technical, logistical, managerial or generative (i.e. the ability to earn legitimacy, to adapt, to create meaning, etc.). They represent a broad range of collective skills that can be both technical and logistical or ‘harder’ (e.g. policy analysis, financial management) and generative or ‘softer’ (e.g. the ability to earn legitimacy, to adapt, to create meaning and identity). All capabilities have aspects that are both hard and soft. (Morgan 2006)

Although the collective nature of capabilities described in the first definition was confirmed by this third definition, some new ideas were introduced. First, as the concept of capability was not related to any specific formally recognized institution but rather to a group or a system, a series of informal attributes associated with the entities articulating and using capabilities as part of their regular functioning were emphasized. Second, although the distinction between “hard” and “soft” capabilities featured by some prior definition of capacity was reiterated, a new concept of “generative” capabilities (capabilities in continued evolution that were characterized by intangible features) pushing for more in-depth evaluation conducted over a longer period of time than traditional assessments, was introduced.

Capabilities (4):

There is always capacity of some degree; capability is how able you are to use it. It’s like I’ve always considered there are clever people and smart people. Clever people may not be able to use their intelligence very well ... in that sense they are

not very smart. Smart people may not be terribly clever but they use what they've got very well. (Contributor to XCeval Online discussion on ECD, 2011)

This fourth definition was particularly useful for the purpose of this study as it explicitly distinguished between capacity (the latent ability of an individual to do something) and capabilities (the actual ability to use one's own capacity in practice). Such distinction became all the more relevant as it appeared to push both ECB and ECD practitioners to support programs aimed at enhancing both capacity (e.g., organizational development which often took a longer period of time to be strengthened) and capability (that is, a set of defined skills of immediate applicability in the interest of the organization in question).

Capabilities (5): "Individual capability is quite distinct from organizational capacity. The difference is all the more important as capacity development has often been reduced to training. Capabilities are a necessary but not sufficient condition for organizational capacity" (former Director of the European Evaluation Society contributing to XCeval Online discussion on ECD, 2011).

An important distinction between capability (regarded as an individual—rather than collective—feature) and capacity (regarded as an organizational feature) was introduced by this fifth definition. As a result, the need for acknowledging the dualism inherent to any activity aimed at supporting individuals and organizations was highlighted. In the same definition, the suggestions also seemed to be made that M&E trainings aimed at enhancing only the technical skills of individuals within an organization (capabilities) would not be sufficient to strengthen the organization itself (capacity)<sup>90</sup>.

---

<sup>90</sup> The limitations of training-only programs had already been highlighted by literature in the past, including in the work of two of the most prolific ECB scholars: "To ensure learning occurs, we must not only be more



### Capabilities (6):

Abilities to do or accomplish something. The distinction between capacities and capabilities is important. A written evaluation policy (capacity) means little if it is not backed by capabilities to commit to and deliver its promises. The application of these capabilities to the three levels at which evaluation capacities need to be established is possible. (Heider, 2011, p. 94)

This sixth and last definition, which drew on Fowler's work (Fowler, 2008, 2011), confirms the distinction between capacities and capabilities presented earlier. It also stressed that interventions enhancing either capacity (e.g., the introduction of new evaluation policies) or greater capability (e.g., the effective readiness to take concrete actions in compliance with the new policies)—but not both of them concurrently—would never be completely successful.

### Additional Remarks on Capability

A term that was often associated with capability in the consulted literature was that of competency, defined as “ [the set of] energies, skills and abilities of individuals” (Morgan 2006). More recently used in debates on the accreditation of evaluation professionals, competencies were increasingly defined as technical skills demonstrated by professionals in a variety of areas (e.g., commission, planning and conduct of evaluation) and aligned with quality standards held up in a more or less formalized way within the evaluation community. The Guidelines on Key Competencies for Development Evaluators, Managers and Commissioners developed by the International Development Evaluation Association (IDEAS) represented a good attestation of this new connotation of the term (IDEAS, 2011).

---

intentional in the design and implementation of evaluations, but we should also complement participation in an evaluation process with other kinds of learning activities and processes” (Preskill & Boyle, 2007, p. 148).

## Capacity Development

Capacity Development (1): “Capacity Development takes place on a continuum between fieldwork, actional learning, and more structured Organizational Development (OD)” (Kaplan, 1999, p. 37).

Although this first definition was not referred directly to the evaluation field, it could easily be applied to it, especially if the commonalities existing between the perspectives of capacity offered by the definition (that is, field implementation and OD are strictly inter-dependent) and the capacity-capability dualism discussed earlier in this chapter were taken into account.

Capacity development (2): “The approaches, strategies, and methodologies used by a developing country, and/or external stakeholders, to improve performance at the individual, organizational, network/sector or broader system level” (CIDA, 2000, p. 12).

Based on the concept of performance usually associated with the RBM paradigm and the more instrumental view of capacity discussed earlier in the chapter, a different view of CD was provided in this definition: no longer regarded as the development of a definite set of skills, CD was associated with the principle of knowhow and described as the readiness to maximize current and future knowledge and skills to enhance performance at several levels (both micro and macro).

Capacity development (3):

Capacity development implies a shift for donors leading to a significantly diminished role in problem identification, design and implementation of interventions and greater emphasis on facilitation, strategic inputs and supporting processes aimed at strengthening developing country capacity. Functionally, this means a move away from ‘donor projects’ to investments in developing country programs and less reliance on expatriate technical assistance. The roles of outsiders, including external support organizations, in such a context, have to be negotiated; they cannot be assumed. (CIDA, 2000 p. 8)

Through the focus on the centrality of national ownership over CD processes, two new concepts were introduced by this third definition. First, a more limited role of donors was envisaged both in the design and implementation of CD programs in the countries that they supported. Second, donors' shift in ECD programming was called for, based on the needs for greater customization and enhanced adaptability of CD programs to local needs.

#### Capacity development (4):

The aim of capacity development is to improve the potential performance of the organization as reflected in its resources and its management. Along with the characteristics of the internal environment and the external environment where the organization operates, capacity influences organization's performance. (Lusthaus et al., 2003, p. 4)

As the functional perspective of CD and the organizational sphere of its applicability (organizational as opposed to individual or systemic) were reiterated, the relevance of the resources and management available within the organization as well as of the contextual factors was emphasized.

#### Capacity development (5):

The process whereby individuals, groups, and organizations enhance their abilities to mobilize and use resources in order to achieve their objectives on a sustainable basis. Efforts to strengthen abilities of individuals, groups, and organizations can comprise a combination of (i) human skills development; (ii) changes in organizations and networks; and (iii) changes in governance/institutional context. (ADB, 2004, p. 14)

This fifth definition resonated with the functional or instrumental view of capacity emphasized by the prior definitions presented in this section: capacity was developed and needed to be developed further to attain some specific objectives in a sustainable manner and not simply on an ad hoc basis. However, one main limitation was identified in relation to this definition. Although the three levels of CD implementation (individuals, groups, and

organizations) had been identified, a clear distinction between groups and organizations was not provided.

Capacity development (6):

Donors and country governments collude to keep the aid game going by using the capacity development label to dress up conventional activities, particularly training. They are, however, not usually able to point too much in terms of real increases in performance. Capacity development thus becomes a kind of symbolic cover, but is devoid of much real content. (Morgan, 2005b, p. 26)

This was by far the most critical definition of CD offered in this section for two main reasons. First, the ineffectiveness of trainings (regarded as the most recurrent activity falling under CD) was condemned. Second, the real value of CD was dismissed and CD itself was qualified as a simple label conveniently attached to donors' agenda in order to facilitate the buy- in of host countries.

Capacity development (7): "The process whereby people, organizations and society as a whole unleash, strengthen, create, adapt and maintain capacity<sup>91</sup> over time" (OECD, 2006, p. 15).

As was the case for the definition of capacity, the meaning attached to capacity development evolved over time and increasingly assumed more holistic features. As a result, this definition set itself apart from all the others presented earlier. First, CD was described more as a process than an activity. Second, the term "individuals" had been replaced by "people," and the term "groups" had been replaced by "organizations," thus acknowledging a) the ecological influences on the success of CD interventions, and b) CD's broader societal repercussions.

---

<sup>91</sup> OECD defines capacity as follows: "The ability of people, organizations and society as a whole to manage their affairs successfully" (OECD, 2006, p. 9).

#### Capacity Development (8):

It encourages debate around deeper meanings of knowledge, learning and change; to better understand the way power relations influence the capacity of individuals and organizations to engage as actors in processes of development and change; and to explore more systemic approaches to learning and change. (Taylor and Clarke, 2008, p. 43)

In this definition the catalytic role of CD vis-à-vis individuals and organizations (mainly in relation to the key goals of learning and change), rather than the procedural aspects of CD, was stressed. As a result, the relational nature of CD was reiterated, as attested by the influence of power dynamics within the interactions established among the actors involved in CD processes.

#### Capacity development (9):

It is about creating conditions that support the appropriate actors in assuming the appropriate roles in this process of identifying problems and defining and realizing goals. More specifically, capacity is derived from establishing effective processes (functions, roles, responsibilities and tasks) for identifying problems or issues, and formulating and realizing goals: these processes are carried out by appropriate actors (individual and collective) who are organized in effective structures for accountability, management and collective voice, and have the motivation, knowledge, skills and resources to perform effectively; the actors are supported in doing so by rules or norms (formal and informal, economic, social and political) that exist within organization (public, private and civil society), in social groups and across society. (UNICEF, 2010, p. 23)

This ninth definition, more than others, pushed for a systemic view of ECD and pointed to the interconnectedness of ECD actors and levels within any given country. As “appropriate actors” were referred to as privileged units of ECD programming within a country system, the need for a value-based choice of who to get involved in an ECD program was also established.

#### Capacity development (10):

[A set of] transformations that empower individuals, leaders, organizations and societies. If something does not lead to change that is generated, guided and sustained by those whom it is meant to benefit, then it cannot be said to have enhanced capacity, even if it has served a valid development purpose. (UNDP, 2009, p. 31)

The description of CD provided by this definition was quite innovative for a variety of reasons. Empowerment and social transformation (Mertens, 2009) (rather than “skills development”) were described as the most critical result of CD. In addition, a renewed sense of agency among CD stakeholders (CD was not “delivered to” but “developed by” stakeholders) was advocated for through this definition. To use a metaphor, the CD scale traditionally tilted in favor of EC providers, whose role was often to transfer capacity (and not always to promote the change virtually associated with it), was now more balanced in the interest of CD service consumers.

Capacity Development (11): “The engine of sustainable human development” (UNDP, 2009, p. 32).

In this rather broad definition, CD was assigned a fundamental role in both the attainment of well-being and the pursuit of effective development at the national level.

#### Evaluation Capacity Building

Evaluation Capacity Building (1): “The intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine” (Leviton, 2001; Stockdill, Baizerman & Compton, 2002).

When this definition first appeared in the special New Directions of Evaluation issue dedicated to ECB in the early 2000s, it proved to be quite innovative for three main reasons. First, organizations and not individuals were represented as the primary ECB stakeholder

group and capacity was alluded to be more of a process than an activity. Second, the intentionality of any activity and program aimed at enhancing capacity was predicated and the underlying assumption that a well-defined plan needed to be in place before ECB could start being implemented, was reiterated. Third, the mainstreaming of good quality and useful evaluation into organizations' day-to-day activities was emphasized as the ECB primary objective. Interestingly, ECB was not always defined as an intentional act. A distinction that started emerging in a number of studies conducted since 2000 was the differentiation of direct and indirect ECB (Bourgeois & Cousins 2008; Cousins, 2003, 2004). Based on this distinction that was originally credited to Stockdill et al. (2002), direct ECB was used to refer to the use of formal and informal training opportunities (workshops and short courses) to build capacity, whereas indirect ECB was used to refer to capacity development "occurring as a consequence of proximity to evaluation, sometimes called process use" (Bourgeois & Cousins 2008, p. 228).

#### Evaluation Capacity Building (2):

A context-dependent, intentional action system of guided processes and practices for bringing about and sustaining a state of affairs in which high-quality program evaluation and its appropriate uses are ordinary and ongoing practices within and/or between one or more organizations/programs/sites." (Stockdill, Baizerman & Compton, 2002)

Some new definitional nuances were introduced by this definition, which drew directly on the principles of intentionality and sustained mainstreaming into organizational functions. First, the relevance of contextual factors was stressed and the need for CB programs to respond to local needs and circumstances was emphasized. Second, given the recognition of the importance of systematic processes and practices, it was suggested that

CB ought to be well structured. Third, based on a multi-level definition of ECB applicability, the inherently systemic nature of ECD goals was highlighted.

#### Evaluation Capacity Building (3):

ECB encompasses a broad range of evaluative tools and approaches that include but go beyond program evaluation. The purpose of the World Bank's ECB efforts is not to build M&E capacities per se; capacity building is simply one step along a "results chain." [...] If a country is able to increase the availability of monitoring and evaluation information and evaluation findings but this evidence is not used, the ECB effort will have failed." (MacKay, 2002)

This definition was quite innovative in that ECB was described as a "means to higher development goals and not an activity aimed at enhancing purely technical functions" (phone interview with Ketih MacKay on February 29, 2012). Interestingly, this definition was not very well supported by in-country practitioners interviewed in the field. When asked about the ECB objective, seventy-three percent (73%) (65/90) indicated that ECB was primarily conducted to enhance the data collection and reporting skills of donor-funded projects staff. An additional fifty-four percent (54%) (49/90) mentioned the ECB primary objective was to improve the project performance and attain the expected project results. In addition, seventy-seven percent (77%) (69/90) of respondents clarified that the topic that ECB emphasized the most in their respective countries was the logical framework, also referred to as logic model<sup>92</sup>.

#### Evaluation Capacity Building (4):

Intentional process to increase individual motivation, knowledge, and skills, and to enhance a group or organization's ability to conduct or use evaluation. (Stockton, Baizerman, and Compton, 2002)

---

<sup>92</sup> This last finding was aligned with a recent systematic review of ECC studies (Labin et al, 2012), according to which ECB appeared to primarily focus on data collection (as confirmed by the 77% of studies) and data analysis (as confirmed by 50% of studies).



The specific ECB objectives were identified by this definition: besides the traditional increase in knowledge and skills, the relevance of enhancing both individual motivation and organizations' ability to use evaluation findings were stressed.

Evaluation Capacity Building (5): It is much wider than training. It involved strengthening or building M&E systems, especially country-based systems, so that M&E is regularly conducted and used by countries and organizations themselves" (World Bank, 2005).

This definition was critical for two main reasons. First, the ECB scope was identified as much larger than commonly thought (ECB was not just about trainings). Second, the instrumental role played by ECB towards the strengthening of countries' M&E system (and not simply towards development of evaluation skills of a few individuals or the evaluation function within a small organization) was emphasized. Interestingly, this definition of capacity building as "more than just training" was not confirmed by the stakeholders interviewed in the field. First, of the in-country practitioners interviewed either over the phone or in person as part of this study (80 in the DRC, Niger and South Africa and 10 in other countries), eighty-eight percent (88%) (79/90) reported that ECB mainly consisted of short-term activities (on average, between 1 day and 2 weeks in duration). In addition, seventy-five percent (75%) (68/90) of respondents indicated that the term ECB evoked to them the idea of trainings and workshops. In addition, forty-seven percent (47%) (42/90) of respondents stated that most of such works were provided by non-national trainers not always selected according to a competitive process. Such findings were confirmed by a recent systematic review of ECB recently published in the American Journal of Evaluation (Labin et al., 2012), according to which ECB was mostly associated with the idea of

trainings, as attested by seventy-seven percent (77%) (47/61) of the articles included in the study<sup>93</sup>. Interestingly, only thirty-two (32%) (37/114) of the respondent tended to appreciate ECB in more comprehensive terms. Interestingly, sixty-two percent (62%) (23/37) of those who considered ECD in broader terms than just had started working on ECB in the early 1990s. Furthermore, twenty-two percent (22%) (8/37) of these respondents had either been involved or were still being involved in ECB programming for larger international organizations

#### Evaluation Capacity Building (6):

Although discussions on evaluation capacity building often focus on strengthening the technical skills of the evaluation practitioners who design and implement evaluations, it is important to appreciate that at list five groups are actively involved in the evaluation process: agencies that commission and fund evaluations, evaluation practitioners, evaluation users, groups affected by the programmes being evaluated, and the public opinion [...] Evaluation capacity building can be delivered in many different ways, formally and informally, through extensive university of training institution programmes or very rapidly. (Bamberger 2006, p. 209)

Thanks to this definition, the multi-faceted nature of ECB (e.g., trainings, workshops, distance learning, on-the-job training) as well as the variety of stakeholders involved in it was identified. However, the encompassing description of ECB attributes and targets included in this definition did not seem to be supported by the majority of stakeholders interviewed in the field. On the one hand, as mentioned earlier, the definition of ECB was associated with the idea of short-term trainings only by the majority of respondents. On the other hand, the definition of ECD targeting provided by in-country practitioners interviewed during fieldwork was a bit more limited than the one provided in Bamberger's definition.

---

<sup>93</sup> Technical assistance (Berg, 1993), reported in sixty-two percent (62%) of cases seems to be more common among those organizations that dispose of more financial resources.

Sixty-seven percent (67%) (60/90) of stakeholders confirmed that ECB mostly targeted national staff working for international NGO and sixty-five percent (65%) (72/90) of respondents believed ECB program targeted national ministries and other government staff. Only thirty-two percent (32%) (28/90) of respondents reported that beneficiaries included representatives from smaller CSOs and the private sector (either working as freelance or in affiliation with private consultants). Similarly, as far as the venue of capacitation was concerned, seventy-two percent (72%) (65/90) of respondents reported that most ECB programs took place in the national or provincial capital and very rarely in rural areas.

#### Evaluation Capacity Building (7):

When we use the term ‘capacity building’, we need to be clear about where we should be focusing our efforts. The term has been used as a euphemism for fundraising, utilizing new technology, increasing and training staff, and even purchasing vehicles. As practitioners who seek to influence positive change in a wide variety of contexts, we suggest that the capacity we most urgently need to build is our own capacity to creatively think and innovate. In order to be able to take the risk of approaching our work in new and unconventional ways, and reflecting and learning from these experiences, we need to create and sustain a safe and supportive environment. (CDRA, 2007; p. 34)

In this definition, which drew upon the concept of “soft capacity” discussed earlier in the chapter (Horton, 2003), capacity was intended not only as enhanced technical expertise but also and foremost as a creative and innovative force that could only be nurtured in a learning-friendly environment.

#### Evaluation Capacity Building (8):

Evaluation capacity building involves the design and implementation of teaching and learning strategies to help individuals, groups, and organizations learn about what constitutes effective, useful, and professional evaluation practice. The ultimate goal of evaluation capacity building is sustainable evaluation practice—

where members continuously ask questions that matter; collect, analyze, and interpret data; and use evaluation findings for decision-making and action. For evaluation practice to be sustainable, organization members must be provided leadership support, incentives, resources, and opportunities to transfer their learning about evaluation to their everyday work. Sustainable evaluation practice also requires the development of systems, processes, policies, and plans that help embed evaluation work into the way the organization accomplishes its strategic mission and goals. (Preskill & Boyle, 2008, p. 149)

This definition was far more encompassing than any earlier articulation of the ECB meaning. However, as ECB was more often than not associated with short-term training and given that ECB was intended to create something that was not there yet, this definition did not seem to take into account the limitations associated with the implementation of ECB activities.

#### Evaluation Capacity Building (9):

Donor centric financing and reporting displaced the purpose of producing global public goods that donors meant to support. [As a result,] ECB is characterized by multiple idiosyncratic projects disconnected from overarching strategic objectives and metrics that disempower leadership as well as by punctual capacity building activities organized around some general ministry strategic objectives. (McAllister et al., 2008, p. 34)

In this definition—one of the most critical provided in this chapter—the disconnect was identified between the limited and often unclear scope of activities funded by donors as part of ECB programs, and their corresponding higher developmental objectives. As a result, the ECB negative repercussions on leadership development and its influence on the perpetuation of an apparent status quo in national development processes were denounced.

### Evaluation Capacity Building (10):

In its narrowest sense, ECB refers to building the skills and ability (human capital) of evaluators around using systematic research methods to evaluate the performance of projects, programs, country development strategies, and global programs. (McAllister, 2011, p. 210)

In this definition an instrumental view of ECB was reiterated and, through the emphasis on the technical attributes of the skills and abilities strengthened by ECB, the use of evaluation capacity for programmatic purposes was highlighted.

### Evaluation Capacity Development

Evaluation Capacity Development (1): “The long-term change process of unleashing, strengthening and maintaining evaluation capacities. It takes place in the context of ongoing efforts to strengthen related systems of management, governance, accountability and learning” (African Development Bank 1998; OECD, 2010).

Two new ideas were introduced through this definition. First, the latent nature of capacity (capacity is not created but unleashed) was highlighted and the ECD “facilitative” and “instrumental” role (rather than the “creational” one inherent to the concept of ECB) was emphasized. Second, through the identification of the systemic nature of ECB goal (“strengthen management, governance, accountability and learning”), the scope of evaluation functions was larger than the one that had been traditionally associated with evaluation in the past (that is, learning and accountability)<sup>94</sup>. Interestingly, the original version of the definition was developed by the African Development Bank in 1998 and expanded by OECD a decade later. As stated by one individual evaluation practitioner in South Africa: “ECD is not an

---

<sup>94</sup> Potter and Brough propose a comprehensive approach to capacity building that incorporates nine elements of capacity. Their approach—illustrated by a pyramid in which the building blocks of systems, staff, skills, and tools support one another through a series of feedback loops—challenges the idea that capacity can be enhanced through the strengthening of a single component. (Baser et al., 2009; Potter et al., 2004)

end it and of itself. ECD is a means to enhance the quality of in-country evaluations so as to attain impact.” In line with this definition, a quite vivid image of ECD was provided by a practitioner actively involved in a regional VOPE in South-east Asia:

ECD is not only about external accountability, that is accountability to donors or funders. It is also and foremost about internal accountability. That is also favored by the fact that lines and communications among ECD stakeholders clearly spell out roles and responsibilities.

Furthermore, as mentioned by one government official in South Africa interviewed in the course of fieldwork:

Unlike ECB that has a more individual-focused approach, ECD is about people and systems. ECD objective is to expand further some ongoing and existing dialogue on how to best use the country M&E systems to enhance development processes. In order for that to happen, it is of utmost importance to have a national evaluation policy in place as well as a large number of evaluation champions<sup>95</sup>.

Although strong evaluation champions were often described key ECD allies, the often-ignored concept of followership was just as important to explain the success of ECD programs. Individuals (referred to in this study as peripheral champions), that is individuals within the systems that were most interested in following the selected primary ECD champions, ought to be targeted by EC initiatives as they are particularly keen in learning, implementing, and disseminating ECD-related knowledge and skills:

---

<sup>95</sup> This last statement introduces the central role played by champions and the relevance of long-term ECD initiatives specifically for the purpose of enhancing “evaluation leadership” within country systems. Champions at all levels are identified within a national system (e.g., private organizations and ministries) and strengthened over time, by minimizing the risks of staff turnover as much as possible. The crucial role of such process is attested by numerous publications, including two large studies conducted by the Global Leadership Initiative (GLI, 2007) and the European Centre for Development policy Management (ECDPM, 2008) respectively, as well as number of governance evaluations. The idea that individual leaders could be instrumental to both bringing about change and fostering innovation is also confirmed by a parallel wave of studies (Uphoff, Esman and Anirudh, 1998). In addition, the work of several international development agencies, has attested the recognition of leadership development processes: UNDP has funded more than 700 leadership courses in 30 countries over the last decade and the World Bank Institute (WBI) has been providing customized trainings on leadership to a large variety of decision makers and managers in developing countries over the last 30 years.

Respondents [...] emphasized the importance of strong leadership for capacity building in the International Union for Conservation of Nature (IUCN) in Asia, but many pointed to an equally important and capable 'followership'. Through collaboration and collective strategic thinking, leadership in the organization is sometimes diffused and shared. (Morgan, 2005b, p. 37)

Certainly the identification of evaluation champions was not to be regarded as a one-time activity but rather a continuous process. Such focus on the processes and the longer time frame associated with ECD programs was confirmed by several of the respondents met during the field data collection. As one former AfrEA president interviewed within the scope of this study put it: “ECD is a gradual, incremental and opportunistic process.” As described by another ECD scholar, ECD was “a fluid and responsive process built upon a system/institutional needs assessment as opposed to the more project-based and logframe-based approach characterizing ECB”<sup>96</sup>.

Evaluation Capacity Development (2): “The ability of an institution to manage information, assess program performance, and respond flexibly to new demands” (Picciotto, 1998, p. 24).

In this definition, in which ECD was primarily described as an institutional attribute, sound information management and performance assessment were identified as the two distinctive signs of ECD articulation. In addition, far from being defined as a static institutional feature, ECDs’ flexible nature and its responsiveness to ever-changing demands were emphasized, as attested by one of the respondents (a seasoned practitioner with over 20 years of experience in a large international organization): “Unlike ECB that tries to

---

<sup>96</sup> Such a statement seemed aligned with a recent statement from Patton, according to which “the CD discourse reiterates two major evaluation trends of the last decade: increased attention to capacity as a critical foundation for useful evaluation and systems thinking as a framework for understanding the dynamics of evaluation use” (Patton 2012. Presentation given at the State Department in June 2012).

disseminate best practices (narrow solutions to clearly defined issues), ECD aims at offering “best fit” practices in an attempt to address unexpected challenges.”

#### Evaluation Capacity Development (3):

Internalization of an effective and efficient M&E practice within the frameworks of public governance systems and institutions that enables tracking of and providing feedback, with empirical evidence, on progress of implementation of public policies and programmes on a timely and regular basis. (Contributor to the XCeval online discussion on ECD, 2011)

Based on the identification of the link existing between ECD and public governance, the accountability purpose that ECD serves vis-à-vis national policies and programs (and not small projects as was the case in earlier characterizations of ECB) was emphasized in this definition. As a result, the idea that ECD was key to promoting national ownership of M&E practices was supported by this definition, as confirmed by seventy-three percent (73%) (65/90) of in-country practitioners interviewed during the fieldwork who mentioned that the creation of an in-country sustainable evaluative culture was one of ECDs’ primary objectives. Such finding was supported by sixty-seven percent (67%) (60/90) of in-country respondents who mentioned that ECD aimed at reducing countries’ dependence/reliance on external technical support. In addition, fifty-nine percent (59%) (53/90) of respondents mentioned that ECD aimed at enhancing the use of evaluation within national governments.

Interestingly, the ECD objective of promoting national ownership over M&E processes was often linked to that of fostering in-country multi-stakeholder planning processes. When asked which groups ECD target population consisted of sixty-seven percent (67%) (60/90) of respondents mentioned a combination of specific ministries, CSO networks and the private sector. Only twenty-nine percent (29%) (26/90) mentioned academia as an ECD target. When asked about venues where ECD took place, fifty-four



percent (54%) (49/90) of respondents reported that ECD activities were normally held within or in the proximity of centralized or decentralized branches of governmental agencies, academia, and CSO networks.

#### Evaluation Capacity Development (4):

[In response to the metaphor that ECD is like planting the tree, nurturing it, irrigating it, protecting it, pruning where necessary], I'd argue that the tree does not develop the apple. That's just the infrastructure on which it grows. Developing an apple is a very complex mixture of atmospheric, physics, biochemistry, genetics, plant physiology, human nurturing. And superstructure is an option – you can probably grow apples these days without a tree! (Contributor to the XCeval online discussion on ECD responding to the definition presented above, 2011)

In this definition, ECD was defined as an endeavor aimed at developing existing capacity further rather than creating evaluation capacity from scratch through a defined mix of activities. Through this definition, the context-responsive and creative nature of ECD programming was emphasized. The creative nature of ECD was confirmed by the majority of respondents who were aware of the term. For the sake of accuracy, of the 64 French-speaking respondents interviewed during data collection, only thirty-nine percent (39%) (25/64) of those who were aware of the term and its use in the English language. Overall, in response to what the term ECD evoked to them, sixty-eight percent (68%) (51/75) of stakeholders who were aware of the term, mentioned the combination of multiple activities: training, mentoring/coaching, technical assistance and peer exchange.

#### Conclusions of Literature Review on ECB and ECD Definitions

Despite the overlapping use of ECB and ECD in some of the literature produced over the last two decades, the distinction between the two terms became apparent only during the last few years, as also attested by the majority of the practitioners interviewed in the field.

That notwithstanding, identifying the differences between ECB and ECD proved to be no easy task, mostly due to the fact that both of them could have some of their delivery modalities and objectives in common.

On the one hand, the definitions of ECB presented earlier in the chapter certainly suggested that there were different modalities of delivery associated with it<sup>97</sup>. Far from being a monolithic concept or being associated with either a positive or negative connotation, the term ECB seemed to derive its quality and effectiveness by the type of goals it was pursuing (the more short-term the least likely to succeed) and the variety of strategies and delivery modalities that came with it (the more integrated the more effective).

On the other hand, ECD appeared to generally be more encompassing than ECB and more focused on bringing about change at the institutional and systemic/societal level than at the individual level only. Differently from ECB, which tended to be more intentional and structured, ECD was also characterized as more adaptive, opportunistic, fluid and responsive to emerging needs.

The fact that the two terms were not always used interchangeably with the same connotation was confirmed by several respondents. When asked whether they believed that ECB was different than ECD, sixty-nine percent (69%) (52/75) of respondents declared that they perceived there was indeed a difference between the two terms. When asked to elaborate a bit more on such difference, seventy-eight percent (78%) (51/75) respondents confirmed that ECB, mostly associated with trainings, was no longer effective or sustainable.

---

<sup>97</sup> ECB strategies have often been assigned a number of strategic objectives to pursue: i) stimulate policy-makers' demand for evaluation; ii) improve the quantity and quality of M&E products; iii) ensure the cost-effectiveness of M&E products; and iv) promote the sustainability of M&E systems and institutional arrangements.

Furthermore, nearly fifty-nine percent (59%) (44/75) of respondents stated that the idea of “building” capacity did not do justice to the fact that there is already capacity in country.

Such dichotomy did not come as a surprise based on the respective etymological definitions. On the one hand, the original meaning of the term building, which originated from the Old English lexicon *byldan* meant to “construct a house,” and its proto-Indo-European variation *\*bhu* was primarily used with the meaning of “to dwell.” On the other hand, the term development, which derived from the Latin term *evolvere*, was already used in the 1640s with the meaning of “to unfold, open out, and expand.” The term, which was composite and resulted from the combination of *ex* (*equivalent to the word "out"*) and *volvere* (*equivalent to “to roll”*), started being used in 1832 to refer to the action of “developing by natural processes to a higher state”.

As the definitions presented earlier in this chapter did not provide specific operational definitions of key ECD concepts, the availability of a more easily understandable and practical framework for action in the ECD arena would be particularly beneficial. However, as the specifics of ECD seemed to be contingent on the needs and specific contexts where evaluation capacity was to be developed, it became apparent that a visual comparison of the main ECD and ECB attributes (as delineated in the course of both the systematic literature review and interviews with ECB practitioners conducted as part of this study), would be particularly beneficial .

Furthermore, when asked about ECB shortcomings, one respondent mentioned the fact that “ECB used evaluation and the development discourse to displace countries and take them off their route to development.” Another respondent in South Africa qualified ECB as an “atomized or drop in the ocean approach in that its activities are generally quite brief in duration, are rarely characterized by follow-up and often involve a limited number of

individuals scattered among agencies not related to each other.” Furthermore, another respondent in Niger, talks about ECD as “a paternalistic support of in-country entities as these were vulnerable and did not possess any minimum level of capacity.” Similarly, a respondent in DRC described ECB as a “unilateral transfer of knowledge, resources and technology.” In addition, an ECD scholar talked about ECB as the expression of “a linear, oversimplified understanding of reality, more worried about testing predictive capacity than capturing the emergence and complexity of unexpected and unknown capacity”

## CHAPTER V

### VOPEs ROLE IN NATIONAL ECD PROGRAMS

#### Introduction

In response to the second research question on the extent to which ECD targeting in international development contexts could become more inclusive in the future (Dahlgren, 2011), both the magnitude of VOPEs' current contributions to ECD programs and the feasibility of VOPEs' further involvement in this area in the future were explored. In particular, three case studies on national VOPEs operating in three different countries in sub-Saharan Africa (Democratic Republic of Congo, Niger and South Africa) were developed. The unique experiences of the three VOPEs presented in this chapter were particularly instrumental in identifying factors that ECD planners and implementers might want to take into account when defining the scope of ECD programs specifically targeting VOPEs in the future.

Based on the case studies, a separate ECD framework was developed for each of the three national VOPEs with two primary objectives. First, to situate VOPEs within their respective countries' national ECD contexts. Second, to highlight the type and quality of interactions existing between VOPEs and other national ECD stakeholders.

## Case Study 1: South Africa Monitoring and Evaluation Association (SAMEA)

### Introduction

South Africa (SA) was regarded as one of the countries—if not the country—with the strongest M&E system in sub-Saharan Africa for a long time. Such widely held belief seemed well justified due to three main factors: a) the prominence assigned to the evaluation function across all levels within the SA government; b) the large number of evaluation specialists and private consulting firms providing their professional evaluation services to both the national government and a variety of SA-based development organizations; and c) the relatively high pressure exerted by civil society on both the public and private sector towards higher accountability and a more evidence-informed allocation of national budget resources to effective development interventions.

Though straightforward, such explanation was perceived to be a bit too simplistic as it failed to capture the multifaceted nature of both evaluation activities and processes in SA. As a result, the need for both a more thorough analysis of the state of evaluation in SA and a better understanding of the extent to which VOPEs and other ECD stakeholders contributed to it was identified. The case study on South Africa presented here—the first of the three included in this chapter—was developed between February and May 2012 to address such needs in two different ways. First, it was intended to describe the role of the South African Monitoring and Evaluation Association (SAMEA) in the creation and strengthening of the SA national M&E system. Second, it was meant to both gauge SAMEA's current capacity and identify its main strengths and weaknesses, especially in view of the possibility of further Association involvement in future ECD programming. Overall, the goal of this case study was to provide ECD planners and implementers with strategic—

albeit exploratory—information on how maximize the effectiveness of ECD programs in countries characterized by long-established VOPEs both at the national and sub-national levels.

This case study consists of four main sections. In the first section, the history of M&E in SA since the mid-1990s was reconstructed and the four main phases of SAMEA history (Pre-formalization, Conceptualization & Formalization, Relational Articulation & Catch-up, and Expansion & Regional Consolidation) were identified. In the second section, the SAMEA organizational capacity (herewith referred to as capability) was described, followed by a detailed description of SAMEA's contribution to the promotion of a national evaluative culture. In the third section, a more general overview of the institutional arrangements adopted by the SA national government to promote the evaluation function at all levels as well as the contributions to it made by other ECD national stakeholders was provided. In the fourth and last section, a few factors that appeared to be either enabling or hindering the SA evaluation efforts were presented.

### SAMEA History

Based on the reconstruction of the most salient events characterizing SAMEA's history since the inception of its activities, three main phases were identified: a) a pre-formalization phase (2002-2004), b) a conceptualization and formalization phase (2004-2005); c) a consolidation and regional expansion phase (2006-2012).

#### First Phase: Pre-formalization (2002-2004)

Although the 1996 South African Constitution declared that the national parliament and provincial legislatures had the mandate to facilitate and promote public participation so that citizens' voices could be heard in the national political arena, it took a few years until

citizens in SA were able to both organize themselves and start contributing to national policy setting. The evaluation professionals working in SA were no exception. Despite their involvement in a variety of evaluation assignments since the 1990s, SA evaluators gathered for the first time only in April 2002 in order to create what would be known a few years later as the South African Monitoring and Evaluation Association (SAMEA)<sup>98</sup>. Interestingly, the event that paved the way for the establishment of a national evaluation community in SA was an M&E workshop delivered by Michael Quinn Patton (an internationally renowned evaluator from the US) in front of an unexpectedly large audience of three hundred and fifty individuals attending the event and sharing with each other a common interest in evaluation.

It was in the aftermath of this milestone event that Zenda Ofir (one of the meeting participants and already a well-established evaluation professional in South Africa) coordinated the idea of creating a new national evaluation association. Prior to this, she had formed an informal community of practice under the name of South Africa Evaluation Network (SAENet) that was supported by her own business, EvalNet. The Network rapidly gathered the support of seventy individuals. In the meantime, several participants who had attended Dr. Patton's workshops and who were serving as civil servants at the time were able to convince the then-SA president (Thabo Mbeki) to establish an M&E office within the President's Office. Such decision was followed by the designation of an M&E Minister under the following president (Zuma).

The following years were characterized by an increasing collaboration between SAENet (led informally by Zenda Ofir since 2002) and the Public Service Commission (PSC) whose evaluation vision was especially inspired by Indran Naidoo (the PCS Deputy

---

<sup>98</sup> In reality, a South African Monitoring and Evaluation Associations (SAMEA) already existed under the leadership of a non-South African. However, its membership was limited to a few individuals.



Director General). The link between the two entities was so close that they even partnered to sponsor the third annual AfrEA conference held in Cape Town in 2004. The event was extremely beneficial in increasing the Network membership, estimated to be around 400 members by the end of 2004.

#### Second Phase: Conceptualization & Formalization (2004-2005)

Following the AfrEA conference in 2004, a plan was developed to create a National Evaluation Association. In order to do that, a highly participatory consultation process was initiated: a Task Team of 16 volunteers was given a mandate to formalize the conceptualization, governance, operations, and key priority tasks of a formal evaluation network. While Zenda Ofir originally coordinated this task team, in late 2004 she handed over leadership to Jennifer Bisgard, a US national with a long established presence in SA. After holding both a public stakeholder workshop and an online needs survey and following the registration of SAMEA as a non-profit organization, the Task Team officially launched SAMEA in November 2005. Of the original task team, four stayed on as board members and Jennifer Bisgard became SAMEA's Founding Chairperson.

As of 2005, over 200 government people began getting involved in SAMEA's activities. As attested to by one of the SAMEA members interviewed in Pretoria, the increase in number of members working for the government was the effect not only of a deliberate and well-targeted marketing strategy but also the results of the efforts made by a well-established SA public figure particularly committed to evaluation:

Attendance of the first SAMEA conference was very government-heavy. That reflected quite well the type of strategy that SAMEA was adopting in order to enhance the development of an evaluation community in South Africa: first, sensitizing government officials and then opening the evaluation community to the rest. Indran Naidoo, Deputy Director General at the Public Service Commission, was a key SAMEA supporter

and played an incommensurably pivotal role. He did a lot of advocacy for evaluation in a large number of national departments. He had the right personality, he had the political clout and he had very strong connections with many leaders within the government. He was someone who knew whom to call at all times.

When asked during the European Evaluation Society Conference in Helsinki (October 2012) about what personal characteristics helped him the most in his endeavor to strengthen the visibility and recognition of evaluation in SA, Indran Naidoo mentioned that his great asset was his “being involved in several communities of practice, task forces, as well as formal and informal groups of professionals working in a variety of sectors in SA.” In addition, he mentioned that the “mutual trust existing between him and all of those working on evaluation within and outside the government” whom he interacted with, was particularly beneficial.

#### Third phase: Expansion and regional consolidation (2006-2012)

Over the years, SAMEA pursued a strategy aimed at both expanding its members and consolidating the quality and added value of its activities and programs. In order to do that, SAMEA first diversified its membership by getting a larger number of non-governmental actors (e.g., academics, government officials, individual consultants, NGO and Non-profit organization staff) involved in the various programs and activities that it sponsored. As of 2007, the Association included three main groups of members of equal size: government officials, consultants and evaluation specialists. That notwithstanding, the renewed emphasis placed by the SA national government on evaluation a few years later (e.g., through the creation of a department exclusively dedicated to the monitoring and evaluation of public performance) contributed to a new spike in the number of government officers within SAMEA membership. As of May 2010, according to a sector analysis

conducted of the 348 active SAMEA members, twenty percent (20%) of them were government employees<sup>99</sup>, twenty-two percent (22%) private consultants<sup>100</sup>; fourteen percent (14%) NGO/CSOs staff, and fifteen percent (15%) academics and others, including donors such as UNICEF, USAID, and the Swiss Agency for Development Cooperation (Mouton, 2010)<sup>101</sup> (Table 10)<sup>102</sup>.

Table 10.

*SAMEA Membership (By Category)*

Sector type	Total
Private sector	11
International	16
Universities	26
Research Institutes/ Councils	26
NPO	47
Gvt	69
Consultancies	75
Unidentifiable	78
<b>TOTAL</b>	<b>348</b>

*Source:* SAMEA, 2010

Besides the diversification of its membership, SAMEA pursued a second strategy during this third phase: the promotion of systematic opportunities for building and sharing of evaluation knowledge among the Association’s members. In order to do so, SAMEA put in place a variety of learning and knowledge sharing initiatives accessible to all its members. Such activities included: a) the organization of a bi-annual conference, b) the scheduling of

<sup>99</sup> The increase in the number of SAMEA members working for the government could be explained by the increasingly prominent role assigned by the SA government to the evaluation function, as attested by the creation of a Performance M&E Department (DPME) and the identification of 12 national outcomes to be pursued by all governmental agencies.

<sup>100</sup> While all of the employees of private consulting firms might be SAMEA members, the association’s directory groups all of them as if they were one individual member. Therefore, the number of private consultants currently participating in SAMEA should be revised upward.

<sup>101</sup> According to the same study, the directory also included 1054 inactive members (SAMEA, 2010)

<sup>102</sup> NPO stands for Non Profit Organizations and Gvt stands for government.

periodic evaluation meetings and workshops among its members (nearly a dozen in the Johannesburg-Pretoria area in 2011, and around 15 in the Durban and Kwa Zulu Natal provinces in 2012); c) the organization of special lectures delivered by world renowned evaluation experts in South Africa; d) the set-up of an official SAMEA website; e) the promotion of online discussions through an official Listserv (SAMEAtalk); f) the use of social media (e.g., the SAMEA LinkedIn group; g) the dissemination of articles on SA evaluation, through the creation of a virtual library or Symposium available online and edited by the Wits Program Evaluation group;<sup>103</sup> and h) the distribution of a quarterly newsletter.

In addition to both the diversification of its membership and the strengthening of its community knowledge building and dissemination functions, SAMEA committed to attaining two other objectives during this third phase. On the one hand, SAMEA aimed to improve its internal governance structure. It was decided during this phase that the SAMEA board would consist of nine members and that, in order for SAMEA's programming to remain uninterrupted over time, only a third of the Board seats would be up for election every three years<sup>104</sup>. In addition, SAMEA started working on the creation of a Technical Secretariat and started planning a series of activities that would allow the Association to play a stronger role within the promotion of evaluation in the region. On the other hand, SAMEA made a special effort to forge collaborations with other institutions in South Africa. Aligned with the Association's interest in gaining more visibility both at the national and regional levels, the SAMEA Chair contributed to the conceptualization of the World Bank-sponsored CLEAR Evaluation Capacity Development Initiative, housed at the University of

---

<sup>103</sup> During the first semester of 2012, SAMEA has been working on the development of six concept papers focused on six key evaluation issues (the first one on program evaluation methods was published in May 2012) as well as a number of more direct and uncomplicated case studies ready to use (posted on the Association's website by the end of 2012: [www.samea.org.za])

<sup>104</sup> Board members were also allowed to leave the Board for a year and then come back a year later.

Witwatersrand (“Wits”) in Johannesburg. Likewise, SAMEA opened the Association to a variety of new institutional members, such as the Office of the SA Presidency as well as other international and national development NGOs.

#### SAMEA Capability Assessment

The findings of the SAMEA capability assessment, based on the measurement of five different types of organizational capability (Capability to commit and act, Capability to generate development results, Capability to relate, Capability to adapt, Capability to integrate), are presented in this section.

##### Capability to Commit and Act

Overall, SAMEA demonstrated a good level of consistency over the years between the variety of activities that it implemented and the fulfillment of original mission and objectives (Table 11). In the successful implementation of its mandate, SAMEA also demonstrated a great sense of autonomy (Ellerman, 2011). Unlike the other two VOPEs presented later in this chapter, none of the SAMEA members interviewed during data collection in South Africa expressed any interest in applying for financial support from international development partners, in spite of the financial challenges faced by the Association. In contrast to other countries where national evaluation associations had prospered thanks to the support of international foundations or a number of UN agencies and where the general socio-economic conditions were better than in rest of sub-Saharan Africa, the evaluators encountered in SA believed that ECD did not have to rely on donors support:

“We are a very young association and we have a limited role in South Africa. That notwithstanding, we never thought of asking donors for assistance.”

Table 11

*Consistency Between SAMEA Mission/Objectives and Activities*

<b>SAMEA MISSION and OBJECTIVES</b>	<b>IMPLEMENTED ACTIVITIES</b>
<p><b>Mission</b></p> <p>To cultivate a vibrant community that supports, guides and strengthens the development of monitoring and evaluation in South Africa as an important discipline, profession and instrument for empowerment and accountability</p>	<p>-</p> <ul style="list-style-type: none"> <li>- Evaluation Advocacy to a variety of entities both within and outside the governmental sphere;</li> <li>- Maximization of technology as a means to foster dialogue and exchanges among members;</li> <li>- Organization of activities aimed at professionalizing evaluation practices among members;</li> <li>- Promotion of association's activities among professionals operating in different sectors</li> </ul>
<p><b>Objectives</b></p> <p>To provide a platform for interaction and information sharing among all those interested in M&amp;E</p> <p>To promote high quality intellectual, ethical and professional standards in M&amp;E.</p> <p>To increase the use of M&amp;E theory and practice</p> <p>To promote the development and adoption of M&amp;E approaches and methods suitable to South African and development context</p> <p>To promote post-graduate education and continuing professional development in the field of M&amp;E.</p> <p>To enhance the profile of South African M&amp;E at national and international level.</p> <p>To help build understanding of international developments and trends in M&amp;E.</p> <p>To be a resource on M&amp;E in South Africa</p>	<ul style="list-style-type: none"> <li>- Set- up of an official SAMEA website;</li> <li>- Online discussions through SAMEAtalk (a list-serve);</li> <li>- Use of social media (e.g., the SAMEA LinkedIn group)</li> <li>- Adherence to the AfrEA Evaluation Guiding Principles;</li> <li>- Dissemination of evaluation standards among members ;</li> <li>- Dissemination of articles on SA evaluation both within and outside the association's membership</li> <li>- Organization of M&amp;E workshops both at the national and sub-national levels;</li> <li>- Organization of lectures delivered by internationally renowned evaluation expert</li> <li>- Dissemination of national articles on evaluation through an online platform edited by the WITS program Evaluation group;</li> <li>- Support to members to get their work published in the African Journal of Evaluation</li> <li>- Support of different academic degrees with a specific M&amp;E focus (e.g., the University of Stellenbosch, Pretoria and Johannesburg)</li> <li>- Assistance in the conceptualization of the Clear Initiative;</li> <li>- Support to members' participation in international conferences (AEA, IDEAS, EES, etc.)</li> <li>- Establishment of close links with development organizations (institutional members);</li> <li>- Establishment of partnerships with other evaluation societies within and outside the region</li> </ul> <p>All of the above</p>

### Capability to Generate Development Results

SAMEA's technical and logistical capability was greatly enhanced by its sufficiently steady sources of income and its reliable human resources. With respect to the amount of financial resources mobilized by the Association, the greatest share of them was originated through the collection of its bi-annual conference participation fees. Over the last two years (2010-2012), through the opening of SAMEA's membership to institutional members, such as the DPME and international NGOs, the amount of the association's revenues (all deposited on a regular South African bank account) increased. As far as SAMEA's human resources are concerned, its Board consisted of 10 members serving a 3-year term and regularly attending to their tasks and responsibilities.

SAMEA's effectiveness was greatly enhanced by the skillful use of the Association's website [[www.samea.org.za](http://www.samea.org.za)] as well as the opportunistic exploitation of other modern technologies (e.g., social media) to both foster the dissemination of evaluation knowledge and promote access to evaluation learning resources among its members. Thanks to its partnership with the Wits Program Evaluation Group, SAMEA also provided its members with the opportunity to attend a Virtual Conference in 2010 and a Virtual Symposium in 2011, linked to the SAMEA Conference organized that year. In promoting better accountability of public programs as well as fostering citizen's empowerment (as per its mission), SAMEA aligned its activities with the SA National Development Strategy. In addition, SAMEA members were often recognized publically for advocating in favor of such important causes as cultural sensitivity in evaluation, social justice, gender sensitive evaluation, and empowerment of the most disadvantaged (Andrews et al., 2006).

## Capability to Relate

Since its inception, SAMEA demonstrated great ability to forge alliances and partnerships with other national evaluation stakeholders. During its second organizational development phase (2004-2006), the Association had very close links with the PSC. Later on, especially during the period between 2010 and 2012, the relationship between SAMEA and the SA Department of Performance Monitoring and Evaluation (DPME) grew stronger. SAMEA's proximity to the national government was not an exception: the Association established interactions with several actors operating outside of the governmental sphere. Such was the case of the advisory role played by SAMEA to the CLEAR Initiative in 2011-2012 and the technical support provided by the SAMEA Board of Director to AfrEA on the occasion of their biannual conferences, especially the one hosted in Cape Town in 2004.

SAMEA's success in developing positive relationships with other national evaluation stakeholders was explained by its high level of social and political legitimacy. The social legitimacy, from which SAMEA could benefit, mostly derived from not only the recognition of its long-established mission to promote greater accountability and transparency within the SA government but also from the popularity of its online initiatives. With respect to its political legitimacy<sup>105</sup>, SAMEA was certainly respected by both the government<sup>106</sup> and civil society and its autonomy and independence (mostly due to the diversity of its membership) was widely acknowledged nationwide.

---

<sup>105</sup> SAMEA is an entity legally recognized by the government: it was registered as a non-profit organization, according to Section 21 of the Companies Act.

<sup>106</sup> SAMEA was always in good standing with national institutions. In 2011-2012, SAMEA was even asked by the Office of the Presidency to assist with the development of a national evaluators certification process.



## Capability to Adapt

SAMEA was able to both adapt in a variety of ways and modify its action plans in response to a continuously evolving national scenario. First, in the aftermath of the Department of Performance Monitoring and Evaluation (DPME) creation, SAMEA started pursuing collaboration opportunities with the national government towards the promotion of better evaluative practices within civil society. Second, by shifting away from its traditional focus on promoting learning within its members, SAMEA started linking with both the national government and the CLEAR initiative as of 2011 in order to foster a stronger demand for evaluation and its use among policy makers. Third, in line with the regional consolidation plan pursued since 2007 and in response to the emerging need for more indigenous literature on evaluation, SAMEA established closer links with the African Evaluation Association and started assisting several of its members in getting their research work published in the African Journal of Evaluation. Fourth, concurrently with the creation of new M&E Units within provincial departments, SAMEA started decentralizing some of its training activities so that it could reach out to interested evaluation professionals working in provinces such as Kwazulu-Natal and Western Cape.<sup>107</sup> Based on a Work Plan that addresses a series of capacity needs highlighted at the decentralized level, SAMEA planned the delivery of 10 workshops in Durban (Kwazulu-Natal Province) in September 2012 and the organization of 10-15 workshops in the Johannesburg-Pretoria area in 2013.

---

<sup>107</sup> In 2012, SAMEA was responsible for the organization of a 2 week-and-a-half workshop in Kwa Zulu Natal whose primary objective was to create a M&E Unit within the Department of Economic Development and Tourism under the auspices of the Office of the Presidency. This unit started collecting district-level data on the poorest households living within the province. SAMEA also developed a Memorandum of Understanding with the Western Cape province. Both M&E Units were created at the provincial Offices of the Premier, also known as MEC (Members of the Executive).

## Capability to Integrate

Overall, SAMEA's programming demonstrated itself to be quite responsive to the vast array of interests and needs emerging among its members since the inception of its activities. This was also facilitated by the plethora of networking opportunities aside from the biannual conferences that the Association was able to provide on a continued basis to all its members, thanks also to the opportunistic use of modern technologies. What was particularly striking about SAMEA was the high level of autonomy and independence (financial and programmatic) claimed by the Association. SAMEA also demonstrated great solidity in reviewing the issue of certification and, instead of attending to the DPME specific request to develop the guidelines of credentialing process for evaluation, it committed to promoting further transparency and accountability within the government rather than focusing on the certification of individual technical competencies. As one SAMEA member interviewed in Pretoria said:

I am against the idea of certifying evaluators. The risk associated with the introduction of a certification process would be the creation of a "professional protectionism" and the subsequent loss of vibrancy within the evaluation community [...] Also, how would certification be feasible given the large number of sectors and methodologies currently in use within the community? What I believe it would be good to do is to develop criteria to "recognize prior learning".

As a result, SAMEA decided to continue to provide its members with a variety of professional development opportunities. The sponsorship of courses delivered to SAMEA membership by international evaluation scholars and specialists (including Ray Rist, Michael Patton, Patricia Rogers, Andy Rowe and Marco Segone) as well as the participation of Jennifer Green, Jim Rugh and Michael Bamberger in the last SAMEA conference, are a clear attestation of that.

## Other ECD Stakeholders in SA and Corresponding Institutional Arrangements

### The DPME

Often viewed as the result of individual consultant's efforts or donor-driven initiatives, the promotion of M&E in South Africa did not benefit from the SA national government's support until the mid-2000s. It was only in 2004, in an effort to catch-up with the individual evaluation practitioners' involvement in the M&E discourse held until then, that the SA government started implementing a variety of interventions aimed at recognizing the centrality of M&E in national planning and management processes. Such a shift was partly explained by the SA government's interest in addressing the gaps in both the management and implementation of government-sponsored programs that had been highlighted in a Ten-Year Review of Government's Performance released in late 2003. In order to avert the Review's pessimistic forecasts on the deterioration of public services across the country, the SA government implemented a few innovative M&E initiatives including a) the dissemination of M&E guidelines—an integral component of the National Monitoring and Evaluation Policy Framework—among all the government's departments nationwide in 2005; and b) the government-wide implementation of a national Outcome Approach (OA) in 2009, shortly after the national political elections. Inspired by a variety of management systems used in other countries, including Canada's Management Accountability Framework, and mainstreamed in the day-to-day activities of all departments through the use of score cards, the OA rapidly became the foundation of the SA national M&E system<sup>108</sup> (Box 1).

---

<sup>108</sup> OA rested on a) all departments' commitment to pursue twelve outcomes—referred to as national outcomes—within the scope of their respective portfolio (see Box 1) and b) the commitment of different

### Box 1: OA and the 12 SA national outcomes

1. Quality basic education
2. A long and healthy life for all South Africans
3. All people in South Africa are and feel safe
4. Decent employment through inclusive economic growth
5. Skilled and capable workforce to support an inclusive growth path
6. An efficient, competitive and responsive economic infrastructure network
7. Vibrant, equitable, sustainable rural communities contributing towards food security for all
8. Sustainable human settlements and improved quality of household life
9. Responsive, accountable, effective and efficient Local Government system
10. Protect and enhance our environmental assets and natural resources
11. Create a better South Africa, a better Africa and a better world
12. An efficient, effective and development oriented public service and an empowered, fair and inclusive citizenship

*Source:* DPME 2012

In order to evaluate the public sector performance in attaining each of the 12 national outcomes, DPME started using the Management Performance Assessment Tool (MPAT)<sup>109</sup>, based on a similar tool used in Canada. The MPAT system enabled public managers to assess their own management practices and, by comparing them to those of colleagues in other sectors, to identify management practice improvements that were likely to enable improvements in service delivery (DPME, 2012). As a result, the demand for ECD services within governments grew exponentially within a short period of time. Besides their interest in learning more about data collection, reporting requirements and the use of evidence to make programmatic decisions, what moved national departments to get more

---

sectorial ministries—competent for each of the 12 outcomes—to sign a corresponding Performance and Delivery Agreement (PDA) with the national SA government.

<sup>109</sup> The MPAT system embodies four specific areas of focus: a) *Employees, Systems & Processes*: Organizational Design, HR Management (Planning; Performance; Discipline; Recruitment; Development; Personnel Admin; Retention); b) *Financial Management*: Financial Management; Asset Management; Compensation of Employees; Revenue Management; Transfer Payments; Supply Chain Management and Budgeting; c) *Governance & Accountability*: Management Structure; Ethics; Accountability; Internal Audit; Risk Management; Stakeholder Management and Organizational Culture; d) *Strategic Management*: Project Management and Delivery of Programmes; Monitoring & Evaluation and Strategic Planning. Presentation delivered by the DPME staff at the African M&E Systems Workshop sponsored by the Government of South and the Clear Initiative and held in March 2012, in Pretoria.

serious about M&E was the concern that the public availability of the data they were collecting (in compliance with “let the sun shine” principle supported by the national government) would possibly compromise their future funding. In response to such demand for M&E services expressed by departments, the SA national government even started organizing quarterly sectorial meetings (Implementation Forums) under the leadership of “outcomes facilitators” appointed at the level of Deputy Director General (Clear, 2012).

In addition to that, the President’s Office formed a fully dedicated team of evaluation experts in January 2010 and appointed both a Minister and Deputy-Minister of Performance Monitoring and Evaluation. It took one more year until a Department of Performance Monitoring and Evaluation (DPME) was created within the Presidency with a total annual portfolio of US\$30 million. Thanks to the inclusion of M&E in all departments’ agenda, DPME gained both great public visibility and political clout within the overall government apparatus quite rapidly<sup>110</sup>.

## Chapter Nine Institutions

The DPME was not the only national evaluation champion in SA. The South Africa ECD context was quite unique in that it was characterized by the active promotion of

---

<sup>110</sup> Overall, the development of an institutional architecture aimed at strengthening the evaluation function nationwide was a gradual process, as attested by some of the most recent policy initiatives. In November 2011, for instance, the DPME introduced a National Evaluation Policy Framework (NEPF). Aimed at enhancing both the use of evaluation in government and its impact on government performance, the content of the new policy was informed by the lessons learned by DPME key staff during the study tours conducted in Mexico, Colombia and Australia in early 2011. Although the NEPF original focus was quite broad and encompassed the evaluations of policies plans, programmes and projects as well as the definition of different possible types of evaluations (Diagnostic Evaluation, Design Evaluation, Implementation Evaluation; Impact Evaluation; Economic Evaluation, and Evaluation Synthesis), it recently became narrower. As a result of the shift in the DPME goals, a National Evaluation Plan indicating a specific set of key evaluations to be conducted over the next few years (10 in 2012/13, 15 for 2013/14 and 20 for 2014/15) was approved by the Cabinet in summer 2012.

evaluation implementation and use among a variety of institutions, normally referred to as Chapter Nine Institutions (see Box 2).<sup>111</sup>

**Box 2.**  
**Chapter Nine Institutions**

1. The Public Service Commission (PSC)
2. The Public Protector (an investigation panel following up on instances of non-compliance)
3. The South African Human Rights Commission (SAHRC)
4. The Commission for the Promotion and Protection of the Rights of Cultural, Religious and Linguistic Communities
5. The Commission for Gender Equality
6. The Auditor-General (it looks at public accounts and monitors the service delivery performance. It also provides “special audits” to governments based on demand)
7. The Independent Electoral Commission (IEC)
8. An Independent Authority to Regulate Broadcasting

Among the Chapter Nine Institutions, the Public Service Commission (PSC) was probably the one that played the most pivotal role in the promotion of an evaluative culture in South Africa, especially in the years preceding the DPME creation<sup>112</sup>. That was particularly due to two unique PSC practices. First, the PSC, funded by the Treasury but reporting to the Ministry of Public Service and yet maintaining its independence of it (Chelimski, 2008) was responsible for ensuring that all government’s annual assessment and performance review

---

<sup>111</sup> The terms refers to the Chapter (Chapter 9) of the South African Constitution providing for the creation of such institutions whose primary objective is to guard democracy and contribute to the promotion of a more distinct accountability culture.

<sup>112</sup> Office of the Public Service Commission (OPSC): The OPSC derives its mandate from sections 195 and 196 of the Constitution, is independent and not part of the Executive and reports directly to Parliament. It is tasked with investigating, monitoring, and evaluating the organization and administration of the public service. This mandate also entails the evaluation of achievements, or lack thereof of government programmes. The PSC also has an obligation to promote values and principles of public administration as set out in the Constitution, throughout the Public Service (e.g. professional ethics, efficient, economic and effective use of resources, impartial, fair and equitable service provision, transparency and accountability). As part of its mandate, the PSC produces regular reports on the State of the Public Service as well as other areas of specific interest to the Government (DPME, 2010).

reports were published and disseminated on a systematic basis, thus suspending the old practice of releasing reports two years behind schedule. Second, PSC mandated that government reports were to be published every quarter rather than once a year. As of late 2011, the agency attained some impressive results.

The PSC represented an “independent watchdog” and was committed to improving the public sector effectiveness through three main tools: a) the “State of the Public Service Report” that measured performance of each national department against the nine values<sup>113</sup> of the public administration spelled out in the national Constitution; b) surprise inspections of policy stations and health clinics; and c) a series of “Thematic Evaluations.”<sup>114</sup> As the PSC experience demonstrated, the promotion of better evaluation practices in SA did not end with the DPME creation. Besides its vigilant oversight of the national public sector performance, the PSC was able to contribute some innovative ideas to the national evaluation debate. In order to enhance the national government’s effectiveness, the PSC made sure to promote the mainstreaming of evaluation in the various departments’ lines of work not as a stand-alone activity but rather as a series of well-defined tasks associated with a specific set of technical competencies and skills spelled out in the Service Agreements signed by each government employee and expressed with terms that all civil servants could more easily relate to. As the Chief of one of the PSC Monitoring Units declared in the course of an interview in Pretoria:

Evaluation is a new concept for many people in the SA government and the definition of who is an evaluator or what an evaluator does inevitable poses

---

<sup>113</sup> The nine values included: a) Ethics (initiatives centered on corruption prevention); b) Efficiency (economic effectiveness); c) Development Orientation; d) Fairness and Equity; e) Responsiveness to needs in policy-making; f) Accountability; g) Transparency; h) Good HR practices; i) Representativeness of their staff.

<sup>114</sup> In order to enhance the understanding of evaluation among its own staff, the PSC also developed a short document (a sort of glossary) on basic M&E concepts that rapidly became a reference guide for a large number of government employees.

some challenges. After all, how do you define who an evaluator is? How do you draw boundaries between those who are entitled to be part of the evaluation community and those who are not? How do you open the community so that more people from the public service get a chance to actively participate in it? Maybe you should change the language and define evaluation-related competencies and skills not as if they were unique or any different than those normally envisaged in the ToR and performance review of any other analytical job. Describing evaluation functions by using a language which people in government are already familiar with would enhance the ownership of the evaluation processes and use among a larger group of professionals, such as planners, organizational development experts, cost-benefit analysts, financial managers and accountants, policy development specialists and policy writing advisers.

Other ECD stakeholders that proved to play a significant role in the promotion of an evaluative culture in SA were the Auditor-General<sup>115</sup> (Box 3) as well as the National Treasury, the Department of Public Service and Administration; the Office of the Public Service Commission.

The SA Public Administration Leadership and Management Academy (PALAMA)

Another key ECD stakeholder in SA was the Public Administration Leadership Management Academy (PALAMA). Serving as the training arm of the SA government, PALAMA had become a key actor in strengthening of national capacity in a variety of sectors in South Africa as of the late-2000s. Originally intended as the Government's Technical Assistance Unit (TAU), PALAMA started rolling out a series of training on M&E (one 1-week M&E training and one longer M&E Training of Trainers) in 2010.

---

<sup>115</sup> The annual reports of government departments need to include, inter alia, audited financial statements and statements of program performance. Section 20(1)(c) of the Public Audit Act (25 of 2004) requires that the Auditor General express an opinion or conclusion on "reported information of the auditees against pre-determined (e.g., performance) objectives," Similar provisions exist in terms of the Municipal Systems Act of 2000 and the Municipal Finance Management Act of 2003 at local level. Given the importance of proper and appropriate expenditure within the Public Sector the reports of the Auditor-General feed into the overall monitoring process and do serve as a key indicator of institutional performance in government. The Auditor General also undertakes discretionary performance audits which are very close to evaluations and the exact relationship between these is yet to be determined (DPME, 2012; CLEAR 2012, p. 151-152).



**Box 3**  
**The Auditor's General Office support of the Public Account Committees**

Through a special institutional cooperation unit (stakeholder management unit) housed at its HQ in Pretoria, AG supported the capacitation of accountants nationwide. In addition, AG assisted the Association of Public Accounts Committees (responsible for the implementation of central audit regulation and provisions nationwide) with the organization of three-day trainings on fiscal responsibility and accountability that were on average aimed at 100 participants working at different levels within the government (e.g., members of parliament, support staff, local government officers). As a result of this capacity building program, three main initiatives were initiated. First, a booklet of best accounting practices in South Africa was published and made available among all government employees. Second, a power point presentation on fiscal responsibility was shown during a road show sponsored by AG and held in all of SA provinces. Third, an evaluation training was delivered to a variety of accountants and government employees under the aegis of the AG office.

Likewise, the AG business unit, in charge of developing the annual organizational strategy and providing oversight on the quarterly production of performance assessments (at both the provincial and national levels), contributed to a better general understanding of the auditing processes as well as the use of the respective findings and recommendations nationwide. This initiative was aimed to facilitate the attainment of the AG objective of 100% clean audits by 2015. The Business Unit pursued five main goals: (i) ensure the simplicity, clarity and relevance of the AG Annual Report; (ii) enhance the visibility of AG leadership among auditees and key stakeholders within the government; (iii) strengthen HR through training in auditing skills and the recruitment of over 100 trainees every year; (iv) Mobilize resources (they were no longer in red); and (v) enhance a Lead by Example model in both risk management and internal control and black empowerment.

Far from being a direct-capacity provider of M&E training, PALAMA was conceived to play a more intermediary role in national ECD, that is, to recruit (and provide guidance to) individuals who could then deliver themselves specialized trainings to a variety of other civil servants in SA. As stated in the course of an interview held in Pretoria with the Coordinator of the PALAMA M&E training, the individual trainers selected to work on the

M&E training needed to demonstrate not only M&E expertise, but also methodological and facilitation skills as well as curriculum development experience. Instead of conducting an individual capacity needs assessment before the start of a training (e.g., through the use of a standardized tool), PALAMA normally customized its trainings based on (i) the content review of the “Annual Professional Development Plan” made available by the Department of Labor for each one of the departments that had requested the training, and (ii) a preliminary in-person meeting with the staff that needed to be capacitated. The module on M&E institutionalization commissioned by the SA Department of Roads & Transport was developed for free according to this two-pronged strategy and served as a pilot for the development of similar trainings in other departments in the following months. Interestingly, PALAMA received support from the German International Development Agency (GIZ) in the form of M&E training curricula development assistance rather than direct funding for training delivery.

Recognizing that training effectiveness was hard to measure, PALAMA relied on the use of three key tools: (i) the training facilitator’s report; (ii) the observations of the PALAMA M&E Specialist observing the training; and (iii) the pre-and post-tests (the same set of questions were usually administered to participants before and after the training). The strategies that PALAMA put in place in order to enhance the effectiveness of their current ECD endeavors included five main activities. First, the involvement of Departmental M&E Directors any time a class was offered to their staff. Second, the involvement of a critical mass of government officials (300 practitioners and 200 managers so far) based on a “vertical slice” targeting (entire departments were targeted, from the chief until the most junior staff, before a training was delivered to them). Third, the creation of a Community of Practice (CoP) among former training participants, through the use of a virtual platform.

Fourth, the delivery of evaluation trainings specifically aimed at evaluation managers (e.g., trainings focused on either the development of evaluation ToR or evaluation management strategies). Fifth, in-house mentoring and coaching provided by PALAMA staff.

#### The M&E Learning Network

Similar to the CoP created by PALAMA, the DPME established the M&E Learning Network, a platform enhancing exchange and mutual learning on evaluation among public servants. Based on the assumption that M&E had not yet become the point of entry for self-reflection and learning in government that one would have expected, the DPME initiated this Network to facilitate the gathering, packaging, and dissemination of M&E information and models to a wide variety of practitioners within the governments. As part of this endeavor, a variety of modern technologies (such as a virtual platform, the DPME website, and e-forums) was used in addition to the more traditional modalities of dissemination and knowledge sharing (e.g., newsletters and seminars).

#### Academia

In order to address the weak human capital in evaluation in SA (according to some respondents “universities in SA seemed to be distant from the government and even lowered their standards as a way to enhance access to education to everyone—especially students from marginalized background- to successfully complete their degrees”), a number of universities started offering specific courses and degrees in evaluation:

- The University of Stellenbosch initiated an online master’s degree diploma in evaluation with a strong focus on education (80 students) but with a rather narrow view of evaluation methodology (students only learned about two or three evaluation

approaches and the curriculum made almost no reference to participatory methodologies, such as appreciative inquiry). Recently, a doctoral program in evaluation was also started at the Center for Research, Evaluation, Science and Technology (CREST) (Box 4);

- The University of Johannesburg (under the leadership of Raymond Basson and Fanie Cloete) started offering classes in evaluation;
- The University of Pretoria (Faculty of Medicine) included a few courses in evaluation curriculum design and evaluation design development in its official program (the Policy Study Unit started offering a Master's degree and Ph.D. program in Evaluation in 2012).
- The University of Cape Town (UCT) initiated a graduate program in evaluation directed by Marc Abrahms.
- The involvement of academic institutions seemed to be one of the most

successful ECD initiatives implemented in SA. As stated by a member of the SAMEA

Board of Directors:

The relevance of providing evaluation training to students and young professionals in evaluation is a commendable effort and is likely to enhance the quality of evaluation practices in South Africa. The strategic role of evaluation training is such that the Department of Education should make some more incentives available for the strengthening of evaluation training at SA universities. This could be done, for example, by granting full-time professors accelerated tenure in case they would be willing to include evaluation topics in the curricula of their courses.

#### **BOX 4. The Role of Universities in ECD Programming in SA**

In 2006, under the auspices of Professor Johann Mouton, a post-graduate Diploma in Monitoring and Evaluation Methods was developed and offered at CREST. The online program, open to anyone working on monitoring, evaluation and implementation of public programs and projects consisted of six modules and entailed the obligation for students to produce an evaluation report on a predetermined program of choice in order for them to complete the degree. The six modules were as follows:

- Module 1: General principles and paradigms of evaluation studies
- Module 2: Clarificatory<sup>116</sup> evaluation
- Module 3: Process evaluation and programme monitoring
- Module 4: Data collection methods for evaluation research
- Module 5: Statistical and qualitative methods for evaluation studies
- Module 6: Impact assessment designs

Two compulsory one-week residential courses were offered in April and September each year so as to allow for in-person exchanges among students and between students and teachers. The sharing of learning among the program participants was ensured by WEBCT, the e-learning platform of Stellenbosch University. From 2006 to 2010, of the 549 students who applied to the program, 206 students were accepted for enrollment

*Source:* Mouton, 2010, p. 159-160

#### ECD Stakeholders Outside the Government Spheres

Although the DPME did not engage with civil society since the beginning of its activities, it soon realized that getting civil society involved in the promotion of a national evaluative culture would be quite beneficial. According to the DPME, there were five main reasons for enhancing the interaction between the SA government and other institutional and non-institutional entities nationwide (DPME, 2012):

---

<sup>116</sup> Clarificatory evaluation is the term used in South Africa to refer to formative evaluation. This type of evaluation aims at assessing whether a programme has been properly conceptualized and to what extent it is reaching the appropriate target population. As a result of this evaluation, it is possible to identify what type of programmatic changes are needed in order to facilitate the fulfillment of the envisaged objectives.

1. Multi-stakeholder consultations were likely to put pressure on national governments to be more effective;
2. The participation of multiple stakeholders could foster a more informed dialogue not only on implementation processes but also on needed improvements: far from only taking government officials' views and interests into account, the involvement of non bureaucratic agents could lead to greater inclusiveness and innovativeness of management and decision-making practices;
3. The promotion of civil society engagement was likely to encourage the professionalization of practitioners and a better articulation by the DPME of its needs for the future supply of evaluation services;
4. Closer links with grass-root organizations were key to facilitating a more timely monitoring of both the achieved results and quality of service delivery;
5. Civil society involvement in evaluation could contribute to building both the consensus around and the support of government-sponsored policies thus increasing the level of public ownership.

Following the establishment of the DPME and the OA institutionalization nationwide, the SA government started implementing a series of activities that could be associated with a phase of “relational articulation.” During this phase, DPME envisaged the need for collaborating not only with organized civil society but also with political parties and businesses for a variety of reasons, including increased accountability of public programs. As stated by a DPME official interviewed in his Pretoria office in 2012:

Engaging with civil society is not only a function of civil service delivery but also of accountability. Getting the civil society on board will enhance the

sustainability and successful implementation of the Outcome Approach [...] Through quality engagement with non-governmental actors, the challenges associated with the delivery of public service will be better diagnosed, collective solutions will be found and a culture of cooperation for service delivery will be promoted.

The SA government's engagement with civil society was certainly an innovative institutional practice. However, as stated by a DPME official in Pretoria, working so closely with nongovernmental actors was not always an easy task due to a variety of reasons:

Departments do not like being monitored. They especially do not feel accountable to CSOs and, as a result, they do not view grass-root organizations as potential allies in their effort to improve public sector performance and effectiveness. A large number of department heads continue to believe that they are accountable to the President only and not to civil society.

#### Citizens' Control

The SA performance monitoring and evaluation (PME) framework rested on three main pillars:

- a) the quarterly reporting from coordinating ministers on the degree of progress compared to the delivery agreement targets;
- b) the systematic assessments of the management capabilities of the government institutions tasked with delivery; and
- c) the provision of citizens' feedback on how they experienced the performance of government, with a special emphasis on key selected frontline service delivery areas.

In acknowledging the different features of the three stakeholder groups delineated above (ministries, other government institutions, general public) the DPME started implementing a four-pronged approach aimed at overcoming all the challenges

experienced in attaining the delivery outcomes (unlike frontline service delivery, it is not about assessing the quality of services). First, the stakeholders that the DPME would be interacting with were all identified through an engagement plan. Second, regardless of the specific stakeholder group (political parties, CSOs and businesses), the SA government planned to hold inception meetings with each of them. The purpose of such meetings was to clarify the OA key features and purposes (including DPME responsibilities) as well as the opportunities for and modalities of stakeholders' further involvement. The only difference was that the meeting with the political parties would be a rather broad one while the meeting with CSOs (that is, representatives of CSO networks or umbrella organizations) would be organized around specific outcomes of interest.

Third, it was decided that the relevance and effectiveness of engagement would be evaluated both by the leading department and the CSOs engaged in the process. Fourth, it was expected that the engagement would be reviewed based on the preliminary findings of such evaluations.

#### Potential Risks of Engaging with CSOs

There were some potential risks associated with such engagement processes (DPME, 2012). First, the government could be found liable of uncooperativeness and dishonesty in case the engagement process was either interrupted or unable to fulfill the envisaged objective. Second, some leading departments could express opposition to such an embracing approach. Third, the quality of interactions between the government and civil society was subject to be affected by dynamics and processes originating outside of the engagement arena.



## Possible Solutions

In order to mitigate the risks presented earlier in this chapter, the SA government could pursue the following actions (DPME, 2012): a) monitoring the type and quality of interactions within the scope of the engagement programming and anticipating any unnecessary confrontation; b) separating the engagement process from other past or concurrent processes; and c) emphasizing the added value that working together with non-institutional entities might contribute to leading departments. The good intentions manifested by the SA government were commendable. However, the first instance of collaboration between the DPME and civil society (the Congress of the SA Trade Unions or COSATU) did not work as successfully as expected. COSATU ended up taking full ownership of the M&E collaborative initiative established with the DPME. The relationship between the government and COSATU at some point became so complicated that the Communication Office (and no longer DPME) took charge of the partnership<sup>117</sup>. As described by one of DPME officers:

One of the lessons learned from this episode was that you need to move quicker and push more when you are trying to exploit opportunities to your own advantage. That means that you should bring all the stakeholders together. In this very case, it would have been ideal to consult the unions and hold familiarization meetings with business coalitions right away rather than waiting for more general consensus to be reached. Once you have created this enabling environment, you could then take the issue to the Ministers.

---

<sup>117</sup> The idea to collaborate with COSATU was not a deliberate DPME choice but rather the attempt to respond to their accusations of DPME being an instrument of a “Predator’s State and not a Workers’ State.

## Two DPME engagement models and the Department of Basic Education's experience

DPME implemented two main engagement models. The first one (centralized model) called for the government to manage the whole dissemination process of information on evaluation (e.g., through the creation of a forum hosted by the DPME) and was pursued with a certain number of departments. However, the risk associated with the first model was to trigger the opposition of different departments as well as a certain level of resistance from the political opposition.

The second engagement model (organic model) consisted in leaving each department to find its own way of complying with the new evaluation policy and consulting with civil society. Such was the case of the Department of Basic Education (DBE), which created a platform engaging its staff and the national community of researchers and evaluators in 2011. The DBE even established a working committee on Research Endeavors aimed at avoiding duplication of efforts by individual units inside the department. Such initiative mainly consisted of meetings between the DBE staff (clarifying their main interest and developing key evaluation questions) and a selected group of researchers (expressing their opinion on both the technical feasibility and costs of the evaluations proposed by the department officials). The Department of Higher Education also set up a similar forum with NGOs that were involved in research as well as with other research institutions. In order to enhance this dialogue with their government counterparts, researchers started identifying and prioritizing research areas and topics based on the interest expressed by the department itself (in 2012, the theme was Literacy and Numeracy).

## Consulting Firms

Other key actors actively involved in the promotion of an evaluative culture in SA are consulting firms. Such is the case of Khulisa Management services (whose New Initiatives Director served as SAMEA's first President and was currently serving as southern Africa representative in the AfrEA Board at the time she was interviewed), as well as Feedback Research Analytics.

## Relevant Contextual Factors Contributing to the Development of a Strong Evaluative Culture/Enhancing ECD in SA

Some of the distinctive features of the SA case that favored the development of a strong evaluative culture included: a) the tax-based nature of the country's national policies (resources used to fund public programs were mostly indigenous and were funded directly by tax-payers, who were increasingly more sophisticated and better educated than their counterparts during the apartheid era; b); the presence of a strong political opposition providing oversight on the current government's public spending; c) the common practice of conducting data quality audits; d) the high use of cell phones or other social media preventing the withholding of critical information on public program effectiveness; e) the dissemination of report findings among a large variety of population strata (in both urban and rural areas) also thanks to the use of targeted media (e.g., radio in local languages); f) the introduction of greater accountability from public institutions to the general public through initiatives such as a Presidential Frontline Service Delivery Monitoring Programme or the Presidential Hotline citizen complaints management system); g) the effective use of a system monitoring public programs' individual performance<sup>118</sup>; h) the effective decentralization of

---

<sup>118</sup> Between June-September 2011, 122 service delivery sites were visited in five provinces: Gauteng, Limpopo, Mpumalanga, Free State and Northern Cape. These unannounced monitoring visits were to selected Home

the M&E functions through the establishment of M&E units within the Office of the Premier<sup>119</sup> at the provincial level and the development of a more integrated and consolidated approach to reporting on performance.

In addition, SA had a stronger incentives culture than other countries in sub-Saharan Africa. As revealed in the course of an interview with one of the heads of the PSC Monitoring Units, the performance of department heads in SA was appraised on a regular basis and measured against the objectives included in the Performance Agreement included in their annual contracts<sup>120</sup>. Interestingly, measuring the performance appraisal in SA was far from being a simple verification of compliance. Put simply, the production of an evaluation report was no longer considered a proxy of good performance or the end goal of the evaluation function as of 2010. A renewed emphasis began to be placed instead on both the quality and the impact that department actions had on poverty reduction. As declared by one PSC official met in Pretoria:

The excuse of not having the time to either evaluate or follow up on the implementation of recommendations included in past evaluation reports is no longer acceptable. It is rather to be seen as a time management issue, in conflict with Results-Based Management (RBM) good practices that should be systematically adopted within the scope of one's day-to-day activities"  
(PSC)

---

Affairs Offices, the South African Social Services Agency (SASSA) offices, Police Stations, Health Facilities, Drivers License Centres and in some provinces schools and courts. The first cabinet report on the findings from the pilot phase (June-September 2011) was released in January 2012. (The Presidency, 2012). As part of the efforts directed at assisting individuals and communities with public service delivery performance challenges, the Presidency also established in 2003 a hotline for telephonic engagements with citizens. Coupled with this, the Presidency also established a capacity for interactions with individuals through email, via a website, social networks (Facebook and Twitter) and direct correspondence. To ensure that matters brought to the attention of the Presidency are resolved, they are generally referred to the relevant government department or agency and are captured on a tracking system (The Presidency, 2012). Similar initiatives exist in some provincial governments and municipalities.

<sup>119</sup> According to Section 85 of the SA Constitution, the DPME was the custodian of Government-wide monitoring and evaluation. In addition to being the central champion for the outcomes approach, it was expected to serve to integrate data for reporting purposes and to establish a framework for M&E and the development of appropriate capacity.

<sup>120</sup> Certainly, the PSC could not prescribe corrective actions to departments but, if the latter did not fully cooperate, they could be reported to Parliament.

Based on the results of the quarterly reports of the departments, the DPME was able to identify the areas where evaluations were needed the most. However, as it was very important for the central government to get departments (both at the national and provincial level) involved in both the selection of areas to evaluate and the formulation of specific evaluation questions to address, the DPME committed to a more inclusive and participatory M&E process. That inevitably entailed that departments take on some additional responsibilities. As stated by a DPME official met in Pretoria:

You could certainly ask departments what they are interested in and help them in the preparation of their evaluation ToR. However, for this process to enhance ownership and foster mutual learning among the involved stakeholders, departments may need to contribute some of their own resources to cover evaluation costs. That said, the risk associated with the involvement of so many different departments and stakeholders is that there will be confusion on who is responsible for disseminating the results of the research and translating them into operations. The way I see things today is that many of the departments just want to comply with the new evaluation policy and are not really willing to be proactive about the conduct and use of evaluation for better development at the national level.

#### Factors Discouraging the Development of a Strong Evaluative Culture/ECD in SA

As clearly identified by the DPME itself (DPME 2012), three main factors were likely to hinder the development of national evaluation processes in the absence of some adequate and timely actions.

First, the gap between the Government M&E and planning processes. The role and responsibilities associated with the formulation of the Medium Term Strategic Framework, the departmental annual performance plans and the five-year departmental strategic plans did not seem to be too clear. In addition, the DPME role did not always appear well defined and the boundaries between its mandate and those of other institutional actors were not

always clear. Second, the risk of inculcating a compliance culture rather than an evaluative one nationwide. The OA dissemination and the creation of new reporting responsibilities for departments in compliance with the new evaluation policy requirements, if not digested by all the institutional actors involved in the process, could be detrimental to the development of any type of ownership over evaluation processes. Third, the lack of adequately targeted ECD initiatives. Given the variety of actors contributing to the development of an evaluation culture nationwide, the need for more sustained ECD programs recognizing the different needs and interests within society (e.g., within government, academia, the private sector and civil society) was identified.

## CHAPTER 6

### CASE STUDY 2: NIGER MONITORING AND EVALUATION NETWORK (ReNSE)

#### Introduction

The case study presented in this section was developed in April 2012 and focused on the experience of the Niger Monitoring and Evaluation Network (*Reseau Nigerien de Suivi et Evaluation* or ReNSE) This section consists of four main parts In the first part, the six most salient phases of ReNSE history (Informal development, Articulation, Transition-to-national-ownership, Self-reflection phase, Implosion phase, and Nominal Revitalization) were identified. In the second part, the ReNSE organizational capability was described, followed by a detailed description of the VOPE's role in the promotion of a national evaluative culture. In the third part, a more general overview of the institutional arrangements adopted by the Niger national government to promote the evaluation function at all levels as well as the contributions of other ECD national stakeholders, was provided. In the fourth and last part, a few factors that appeared to be either enabling or hindering ECD in Niger were presented.

#### ReNSE History

Based on the reconstruction of the most salient events characterizing ReNSE's history since the inception of its activities, the six main phases of the VOPE's history were identified.

### First Phase: Informal Development (1999-2003)

Although the Niger Ministry of Planning had been engaged in a series of evaluations throughout the 1990s, the level of awareness of evaluation functions and processes within the national government remained quite limited until the early 2000s. Likewise, no evaluation community existed in Niger at that time. Such relatively bleak scenario started changing in the aftermath of a regional seminar on M&E Capacity Development, organized jointly by the World Bank and the African Development Bank, in Abidjan on November 16-19, 1998. The event, which recognized and promoted the role of civil society in the promotion of good M&E practices across Africa, triggered some interest in the creation of a national evaluation association among a number of practitioners, government officers and development partners based in Niger. In particular, it was the UNICEF M&E Officer based in Niger's capital (Niamey) who, building upon his interactions with a variety of in-country stakeholders, paved the way for the creation of the first Niger Monitoring and Evaluation Network (Reseau Nigerien de Suivi et Evaluation or ReNSE) in August 1999. During the first two years of activities (1999-2001), ReNSE particularly benefited from UNICEF technical guidance and financial support<sup>121</sup>. As of 2002, ReNSE started receiving funding from UNDP. Throughout phase, ReNSE was coordinated by the UNICEF M&E officer. In addition, various network meetings took place in different venues from time to time (meeting space was made available on a rotational basis by several UN agencies).<sup>122</sup>

---

<sup>121</sup> UNICEF support during this phase had two primary objectives. First, the promotion of a national dialogue on evaluation in the short-term. Second, the strengthening of national technical evaluation requirements among those who conducted and commissioned evaluations, in the medium- and long-term.

<sup>122</sup> As stated by a member of the UN Thematic Group encountered in Niamey during their monthly meeting, making an "itinerating desk" (that is, an office space changing venue from time to time) available is a good way to keep pushing for the promotion of evaluation among a variety of stakeholders at the national level.



### Second Phase: Articulation (2004-2006)

Although UNICEF had provided most of the funding during the first four years of ReNSE history, a different UN agency (UNDP) started providing the VOPE with financial support as of 2004. The rationale for UNDP support was that the increase in evaluation knowledge and the promotion of a better understanding of both evaluation approaches and methodologies at the national level -which ReNSE was expected to facilitate- would represent a powerful tool to promote better governance and enhanced knowledge sharing (two critical components of the agency's country strategy) nationwide.

### Third Phase: Transition-to-National Ownership (2007-2009)

Building upon the efforts made by the development partners who supported the promotion of a national evaluation culture in Niger during first six years of the ReNSE existence (e.g., UNICEF and UNDP), the Niger government was able to implement a few relevant initiatives during this third phase. First, it established a General Directorate for M&E within the Ministry of Economics and Finance). Second, it created a Ministry of Evaluation in 2007 (this entity lasted only for three months). Third, it drafted a national evaluation policy (the draft was validated in the course of a technical workshop and was being discussed for adoption by the national Parliament as of October 2012).

Furthermore, capitalizing on the success of the third AfrEA general conference held in Niamey in January 2007 and in response to the departure of the UNICEF officer who had coordinated ReNSE activities during the first development and articulation phases, a very active network member took on a coordinating role. The new coordinator, a female national who had led the ReNSE poverty thematic group in the past and who would be nominated Minister of Communication two years later, contributed to the proliferation of the network

thematic groups. However, the renewed focus on the creation of groups sharing a common sectorial interest in evaluation unexpectedly led to a gradual deterioration of individual members' involvement (as also attested by the reduction of exchanges occurring through the ReNSE website). Interestingly, while individual membership dropped quite dramatically during this phase, the number of institutional members grew quite rapidly<sup>123</sup>.

#### Fourth Phase: Self-reflection (2009)

At a time when the network's momentum seemed to have dropped and some effective strategy was therefore needed to rehabilitate the legitimacy of the network at the national level, a discussion started among ReNSE members on the possibility of formalizing the network. After a long internal debate, the ReNSE General Assembly opted for the formalization and rapidly approved the network statute and internal provisions in July 2009 (the formalization was approved by Decree on August 30, 2009). Since the formalization had taken effect, the ReNSE website was made operational again and its management was delegated to a volunteer communication officer (whose continued effort for a period of over two years made it possible to both update the site's content and attend to all other administrative matters quite efficiently throughout this fourth phase). In addition, an independent entity was recruited on ad hoc basis to assist ReNSE with any technical issues (e.g., hardware or server-related issues).

As a result of the formalization, annual fees were also established: as of late 2009, every member was required to pay a one-time initiation fee of 5,000 CFA and an annual membership fee of 10,000 CFA. Despite the optimistic forecasting, the number of paying

---

<sup>123</sup> New members included the Adobu Moumouni University, the National School of Law and Administration, (ENAM), the Institution of Evaluation Strategy (ISEP), GTZ (today GIZ), and CARE International.

members (that is, the number of members holding voting rights) was quite limited during this phase (only 25 per cent of the original network members had paid their fees and had been able to keep their voting rights by early 2010). Although the stream of revenues derived from the payment of membership fees fell below the original expectation, ReNSE was able to mobilize some additional financial resources by receiving from AfrEA a share of the revenue generated through the conference held in Niamey in 2007. In the aftermath of formalization, the network was also able to secure a new office: no longer dependent on donors' availability to share their conference room on a rotational basis, ReNSE found its new home in a two-room office located within the Niamey stadium's premises.

#### Fifth Phase: Implosion (2010-2011)

Following the drop in number of active individual members in 2010, ReNSE started experiencing a few organizational development challenges: its activities became more sporadic and the exchanges over the network's website were interrupted due to the absence of someone within the network who was available to manage it. Well aware of the impasse that the network was experiencing during this phase and in an effort to leave this bleak phase behind, the new ReNSE coordinator decided to react and, as a way to address the identified organizational weaknesses, launched a revitalization workshop in March 2011 (*redynamisation du partenariat* in French) that prompted all members to suggest corrective and feasible solutions to the existing challenges.

#### Sixth Phase: Nominal Revitalization (November 2011- Today)

After a programmatic slowdown, and following the interruption of the online exchanges among the network members, ReNSE embarked on a revitalization phase. In addition to a three-day-workshop on the role of civil society in promoting a national

evaluation culture (November 2011), ReNSE sponsored a variety of activities: a) it held a General Assembly on April 14 2012; b) it sponsored the first edition of the “Niger Evaluation Days” initiative in May 2012; and c) it organized, along with UNICEF support, a week-long workshop on performance and impact evaluation given by a well-respected evaluation professor from Canada (Brad Cousins) in June 2012. The latest ReNSE hyperactivity and the organization of multiple activities generously supported by development partners (including UNICEF and UNDP) was definitely encouraging, but it could not be regarded as a sufficient predictor of effective programmatic recovery or membership reactivation. As a result, it was concluded that the characterization of this last phase, as presented in this chapter, was temporary in nature and that it would take at least six more months for this phase to be qualified more effectively.

As of April 2012 (the last ReNSE General Assembly was held on April 14), the steps planned by ReNSE for the future included the following: a) increasing the Network’s programming focus on planning both at the central and decentralized level (thanks also to the favorable political context); b) fostering demand for evaluation at several levels within and outside the national government; c) providing evaluation-related services; d) certifying (“*normalization*”) the quality of the evaluations conducted at the national level.; e) contributing to the finalization of the text of the national evaluation policy and other relevant strategies (such the 3-year rural development strategy or the 10-year education development strategy); e) enhancing members’ technical capacity (more advanced training in statistics and modeling), especially among the forty percent (40%) of its current memberships made up of agronomists, sociologists, geographers and health specialists; and f) revitalizing thematic groups, such as: Institutional Evaluation; Norms and Methodologies; and Local/Decentralized evaluation.

Through the use of meta-evaluations whose findings would be used to grade evaluation consulting firms, ReNSE was also planning to create positive competition among national evaluation firms and to inform the public administration's recruitment of more qualified service providers

### ReNSE Capability Assessment

The findings of the ReNSE Capability assessment, based on the measurement of five different types of organizational capability (Capability to commit and act, Capability to generate development results, Capability to relate, Capability to adapt, Capability to integrate), are presented in this section.

#### Capability to Commit and Act

Overall, ReNSE demonstrated a fair Capability to mobilize energy and take action. The dedication of a volunteer coordinator and communication specialist was definitely key to the revitalization of the Network. However, the VOPE's internal processes, mainly related to decision-making, did not always appear to be aligned with the statutory provisions, thus calling into question the network's legitimacy among some of its own members. Despite the absence of a formal work plan at the time this case study was developed, ReNSE did not seem to lack planning capacity. The reason for not developing a structured work plan was simply due to the one-time nature of the trainings organized and the limited number of the activities initiated by the Network. This was the case with the formalization process that exhausted most of the network's energies in 2010, the "Days of reflection on the role of civil society in the promotion of an evaluative culture" in 2011, and the organization of an impact evaluation workshop in June 2012. Likewise, every time workshops were planned, a one-

paper (including some approximate date and indicators of progress rather than a well-structured logic model) was developed and shared with development partners (e.g., UNICEF) for possible funding<sup>124</sup>. That notwithstanding, well aware of the relevance of good planning to sound programming and organizational development, the ReNSE coordinating bureau expressed interest in organizing a planning workshop for all its members. As stated by the network Coordinator during the visit to the office in Niamey, the idea for such workshop was “to turn this planning exercise into an opportunity for learning both within and outside the Network and to exploit a temporary weakness as an entry point for strengthening the network’s Capability thanks to the support of other development partners.”

The availability of an office, a Statute and a detailed list of internal provisions, would generally provide adequate basis for the further development of organizational capacity. However, this did not occur in the case of ReNSE due to the fact that some of its organizational practices were not always orthodox or fully compliant with the established rules. The scheduling of the network meetings was a good illustration of that: organized at hours that were incompatible with the busy schedule of most members, meetings generally were not attended by a large number of members, thus calling into question the legitimacy of the network’s decision-making processes. Interestingly, the fact that the decisions were often taken by a very small number of individuals was not questioned by several of the members interviewed in Niamey, given that the “principle of delegation” and “the lack of active

---

<sup>124</sup> Generally, ReNSE contacted donors before the ToRs for any specific planned activity was finalized. As stated by the ReNSE coordinator “sharing ideas with possible funders beforehand” was better than “getting a full-fledged proposal rejected.” Thanks to such approach, defined as a “preemptive approach” (*approche préalable* in the French language), ReNSE was able to secure funding from development partners on numerous occasions.

mobilization” (as it was explained in the course of several interviews) were “very common and accepted within ReNSE.” Another instance that cast some doubt over the democratic nature of the network’s internal processes was the supremacy claimed by founding members over all the other members (every time decisions needed to be made or new initiatives needed to be planned) along with the lack of specific provisions to penalize instances of wrongdoing. The formalization surely did not help dissipate such concerns: some members lamented the fact, for instance, that in order for the voting rights of members to be recognized officially in the course of the General Assembly (April 14, 2012), people needed to pay their membership fees, thus creating a division between what some members described as “first and second class members.” In addition, some members pointed to the unwritten, and therefore questionable, rule that one needed to have served within ReNSE for some time before being eligible in order to become a member of the coordinating bureau. Besides the question of legitimacy of such a provision, its more pragmatic effect is to prevent the network’s coordination from being informed by innovative and fresh ideas.

Overall, the Capability to take action was greatly facilitated by the Network’s ability to mobilize resources both from its own members and other in-country development partners. As of late, membership fees remained the steadiest source of income. That said, a number of ReNSE members who were interviewed, suggested four fundraising strategies that the Network could pursue in what they considered to be a “yet unexplored and virgin territory”:

- a) To conduct performance evaluations in response to Request for Proposals (RFP) issued by public institutions. The Ministry of Planning seemed to be the most suitable governmental entity to support this idea if it weren’t for the lack of specific evaluation skills among the Network’s members. In addition, as the M&E Unit in the

President's Office was already involved in conducting the evaluation of government programs, there appeared to be a need for a closer coordination between ReNSE and governmental actors specifically committed to M&E;

- b) To play a more pivotal role in ensuring the utilization of past evaluations' findings during the development of national sectorial strategic plans. The findings of the Poverty Reduction Strategy Paper (PRSP) evaluation, for instance, were not capitalized upon as they should have been and national strategies (e.g., the Sustainable Development Strategy for Inclusive Development formulated right after the coup d'état) were not always grounded on prior evidence;
- c) To contact private enterprises and submit requests for funding to them that could be eligible according to their respective Corporate Social Responsibility programs;
- d) To become a training provider. Far from being a collective of individual consultants that were exclusively after their own their profit, the network—as argued by many of the interviewed members—could start providing evaluation and capability strengthening services with the caveat that a fixed quota of the possible revenues would be allocated to the Network as a whole. In particular, members suggested that the Network start delivering capacity building activities to both donors and civil society and that it even provide certificates of attendance to those attending its training.<sup>125</sup> According to the VOPE members interviewed in Niamey, the ReNSE's further involvement in this area would be particularly relevant given the relatively low capacity of private consultants delivering evaluation trainings in Niger. The risk associated with such recommended strategy was that only a few qualified individuals

---

<sup>125</sup> When asked about the topics that such training should focus on, members pointed to the following three: a) Monitoring and Evaluation (e.g. of decentralization and democratization programs); b) Evaluation Principles and Standards; and c) Theoretical Approaches



would benefit from the job opportunities becoming available to the Network.

Therefore, for this option to be viable, it appeared that it would first be necessary to provide members with key competencies, including quantitative and qualitative data analysis skills. As stated by some of the members interviewed during a group interview in Niamey:

The idea is not to make ReNSE members become excellent statisticians, but rather to provide them with common terminology and key concepts. Such type of training will need to be very practical, that is, it should be combine theoretical and experiential learning. In addition, as members will not be able to follow a 2-3 day evaluation course, due to their current professional engagement, such training should be sequenced over time and possibly delivered on Saturday and Sunday, without the distribution of any per diem.

Although respondents did not comment specifically on the current network's Coordinator, they seemed to generally appreciate his work, mostly because of his independence, time availability and diplomatic skills. However, some of the members warned that the prior evaluation experience possessed by the ReNSE Coordinator coupled with his privileged access to information regarding professional opportunities in evaluation might give him a competitive edge over all the other Network members.

#### Capability to Generate Development Results

Overall, thanks to the good intentions of both the Network's Coordinator (an independent consultant) and the Communication Specialist (a very motivated young graduate with a degree in evaluation received by a private training institute in Niamey) working for the Network on a voluntary basis, and even more thanks to the diversity of its membership, ReNSE demonstrated a good capability to generate development results. The lack of sustained financial resources (only twenty-five percent (25%) of members were paying the membership fees at the time of the interview) inevitably had an impact on the Network's

technical and logistical capacity<sup>126</sup>. That notwithstanding, the ReNSE provided significant support to the National Government on evaluation-related issues on several occasions.

First, ReNSE contributed to the review of the National Poverty Reduction Strategy 2008-2012 and was part of a Mid-Term Review Committee in 2010. Second, ReNSE served as the Vice President of the M&E Group working on the development of the social and economic development Plan 2012-2015. Building on its prior involvement in national planning processes, ReNSE members suggested that the Network get involved further by doing the following: a) providing quality control on the evaluation conducted on behalf of and committed by the Public Administration; b) offering continued education programs to civil servants on a variety of topics, such as M&E, planning and RBM); c) assisting Research and Planning Units operating within Ministries to put in place training programs; and d) reaching out to young graduates with evaluation skills as to promote the development of a Community of Practice among them<sup>127</sup>.

#### Capability to Relate

At the time when the case study was developed, ReNSE was involved in a series of collaborations with a variety of institutions: the National Community of Practice on Management for Development Results (CoP-MfDR); the International Organization for Cooperation in Evaluation (IOCE), the Francophone Evaluation Network (*Reseau*

---

<sup>126</sup> Interestingly, the ReNSE coordinator was quite realistic in his diagnosis of the network's capacity and expressed great interest in reestablishing a sound relationship with all those development partners who used to support ReNSE in the past, such as the a) the evaluation offices within the UN system; and b) the National Institute of Statistics (the General Assembly meeting was hosted in their building).

<sup>127</sup> One of the ISPE alum created a Support Group (*Groupe d'appui pour la pratique de l'évaluation pour un développement durable*) with other alumni from the school with the idea of supporting the decentralized government entities in their M&E efforts. Unfortunately, due to the lack of funding, such project never materialized

*Francophone d'Evaluation* or RFE in French); the French Evaluation Society (SFE); and the African Evaluation Association (AfrEA) of which ReNSE was a founding member.

Rather than formalizing its ties with other organizations, ReNSE expressed interest in interacting with a variety of stakeholders on an informal basis. Although the Network recognized the value of liaising with civil society, several of its members feared that the Network would lose its neutrality if it espoused social causes advocated by CSO that were criticizing government programs. As a result, ReNSE members were more favorable to foster stronger ties with governmental actors either commissioning or using evaluation, such as the General Accounting Office and the National Assembly. As stated by one of the ReNSE founding members:

The reason for not being part of coalitions with other entities is clear: ReNSE members are very professional and much more experienced in evaluation than the rest of civil society. Therefore, one of the reasons for not formalizing alliances with other CSO is the fear of losing effectiveness.

Other institutions, which several ReNSE members suggested that the Network establish a partnership with, included: a) the international NGOs working in Niger; b) the local municipalities; and c) the private sector (with the scope of Public-Private-Partnerships).

ReNSE appeared to benefit from a discrete level of social legitimacy at the time the case study was developed. Regarded as a group of professionals with specific competencies, ReNSE—as indicated by its Coordinator—was approached by several CSO for support on a number of occasions. Similarly, ReNSE seemed to have a certain political legitimacy at the national level. However, as mentioned by some of the Network's members, all such visibility was not always positive:

The formalization of ReNSE undoubtedly enhanced the Network's visibility and legitimacy. However, it also created huge expectations, such as that ReNSE would be able to

strengthen evaluation skills nationwide as well as to generate new competencies within the national ministries, the Directorate of Evaluation and an increasing number of decentralized entities.

#### Capability to Adapt

Overall, ReNSE demonstrated a fair capability to adapt. Thanks to the intellectual caliber of its members and the close relationship (if not overlap) between them and national policy makers and trends setters, ReNSE was able to serve as a platform to discuss and promote change (e.g., through its various thematic groups). Unlike CSOs largely depending on external funding and therefore not so flexible to respond to emerging needs for which no resource would be made available by development partners, ReNSE was particularly receptive to change. Unfortunately, the formalization of the network and the drop in the number of active members resulting from it, limited the richness and variety of exchanges among members. That notwithstanding, ReNSE remained particularly determined to contribute to Niger's development process. ReNSE's push for the development of a National Evaluation Policy was a clear illustration of that (the ReNSE Coordinator was one of the consultants working on the development of the policy).

#### Capability to Integrate

ReNSE demonstrated a fair capability to develop short-term strategies but a less-than-optimal capability to produce long-term strategies. Also, following the formalization, which in and of itself represented an important organizational change for the Network, ReNSE demonstrated a special appreciation for the new rules and provisions guiding its work. However, in doing so, it seemed to lack (at least in 2011) a sufficient level of flexibility to let, for example, non-paying members participate in the Network's activities.

## Other ECD Stakeholders in Niger

### Capacity Building Providers in Niger I: International Development Partners

*GIZ (German Development Agency)*: serving as the technical assistance arm of the German Development Cooperation in Niger, GIZ had worked to enhance national capacity with several institutional and non-institutional entities in a variety of domains in Niger since the 2000s. One of the most important capacity development endeavors funded by GIZ in-country was the support that had been provided to the national government in both the conceptualization and implementation of the National Poverty Reduction Strategy. Through a series of partnerships established with the Niger Ministry of Planning and other institutional actors, GIZ had been able to disseminate a number of relevant planning and evaluation tools and methodologies. In its effort to foster national ownership of the overall evaluation process, the GIZ (an institutional member) had also been an active member of ReNSE.

The GIZ evaluation capacity development strategy evolved over time. It initially consisted of the organization and provision of individual M&E workshops provided by locally recruited specialists to local GIZ staff as well as key government agencies and civil society organizations implementing projects funded by the German government. It recently evolved, allowing for the seconding of expatriate technical specialists to national ministries as follows:

- a nutrition early warning system specialist was recently placed in the Prime Minister's office for a period of two years;
- one long-term consultant was made available to the Ministry of Planning;

- one more technical specialist seconded to the Office of Subsistence Farming to put in place an Information Management System.

#### UNDP

Until 2006, UNDP used to have a Regional M&E Unit through which it was able to offer M&E support to a variety of stakeholders in Niger. After 2006, the agency interrupted the provision of its capacity building and technical assistance services and started supporting a few activities (including the creation of the National Institute of Evaluation and Perspectives, the first post-graduate program in evaluation available in Niger).

#### UNICEF

A ReNSE partner for quite a long time, UNICEF continued supporting ReNSE indirectly over the years (e.g., through the provision of technical assistance in the development of the National Evaluation Strategy).

#### UNFPA

This agency decided to strengthen capacity at the government level. Since 2004, the amount of financial support for the development of the National and Local Statistical System was equal to \$873,787 and covered four main areas of activities: a) the implementation and analysis MICS Surveys; b) the conduct of the National Census; c) the set up of data collection mechanisms associated with gender-based violence programs; and d) database development.

A Mid-Term Review of the UNFPA program taking place at the time the case study was developed was being conducted by a team of three consultants, recruited directly by the Directorate of Evaluation based at the Ministry of Planning and Finance and responsible for

the coordination of the UNFPA-Niger Partnership Program. While the team was tasked with the data collection, the data analysis was conducted directly the UNFPA office. When asked about the challenges of their M&E work, the UNFPA M&E Officer mentioned the excessive workload and the difficulty of finding suitable consultants who could do the work:

Outside of UNFPA, it is very difficult to recruit evaluation consultants or an evaluation cabinet. Often, UNFPA cannot even finalize the tender due to the lack of a sufficient number of proposals submitted to them. Since 2006, we have been using the same three consultants as they seem to be the only ones who could carry out the evaluations we commission. We are not really sure why such few people are responding to our request for proposals. It might be due to the specifics of the evaluations that we commission.

Similarly, as stated by the UNFPA M&E officer met in Niamey:

We had a couple of very qualified content specialists (a demographer, a doctor and a gender specialist) who did an evaluation for us but they were not methodologists. They certainly did their best in carrying out the envisaged tasks (it was a terrific learning opportunity for them) but they were not evaluators. As a result, some of their conclusions were not based on any findings and often the inclusion of many graphs on trends and regression was a panacea for the lack of a real understanding of the evaluation questions and the issues at stake.

#### FAO

The agency ceased using project-based M&E officers (both at the central and decentralized office) and one of their program coordinators has been seconded by the government to them. However, none of them had received formal training in evaluation. Their capacity building activities included: a) the training of the Planning and Research Unit (DEP) staff at the Ministry of Agriculture; b) the training of the Niger country office and project implementing staff by the M&E expert sent by the Accra regional bureau; c) the delivery of a M&E courses for civil society on new methodologies (e.g., evaluation of beneficiaries' appreciation vis-à-vis the services rendered as part of a food security program) in 2011; and d) the provision of more intense capacity building on sex disaggregation.

Despite the common practice of developing a logical framework for every single project funded by the agency, the FAO M&E officers interviewed in Niamey stated that they would feel more comfortable if they had as a reference a generic CPF (Country Programme Framework). The FAO officers interviewed in Niamey also pointed to the need for better evaluation management technique: “When we go to the field, we do not always use a checklist. What we look at depends exclusively on what we are interested in.”

Other UN agencies involved in M&E capacity building work were identified. These included a) ONUSIDA, which was working quite closely with the Inter-ministerial Committee; b) UNWomen, which supported a series of capacity building programs; and c) IFAD (International Fund for Agriculture Development), which was only active in the Maradi region.

#### World Bank/African Development Bank/3ie

A five-day training on impact evaluation, sponsored by the World Bank and involving a few evaluation practitioners from Niger, was organized in Addis Ababa in 2008. It was the first time for many of the participants to learn about relevant concepts such as baseline and counterfactual. The definition of impact itself was clarified for the French-speakers (contrary to the English speakers, most evaluation specialists in French-speaking countries associate “impact” with the concept of “goal” and not of “effects”). The training then provided a platform to reiterate the constitutive elements of impact evaluation, distinguishing it from performance evaluation or other types of studies and assessments. Learning about experimental and quasi-experimental designs was also particularly appreciated. Training follow-ups do not seem to take place as often as participants wish they did. However, the Internet had been increasingly facilitating post-training peer-to-peer



communication through e-mail exchanges. Most participants were able to share the training content with their colleagues as well as to get their respective directors to buy into the whole notion of impact evaluation. Each research unit within a ministry has a statistician and this function is often regarded as the most appropriate to foster a culture of evaluation along with the unit's director.

The National Institute of Statistics in Niger ([stat-niger.org](http://stat-niger.org)) also received support and seemed to be well positioned to conduct impact evaluation studies. That having been said, randomization appeared to be very problematic (as a recent evaluation of the impact of investment on rural poverty illustrated quite clearly) and a perception survey was believed to be a more appropriate tool to gauge the impact of a development program (Fleischer et al., 2011). Some other times it was the lack of resources that prevented ministries from conducting an impact evaluation (quite recently). Also a ToR was developed to analyze the impact of programs funded by both the WB and UN systems but it fell through in the end as no budget was available for this activity.

A follow-up impact evaluation training was organized in Dakar in 2010. That time, each participating country needed to present a national program that had been identified as the object of an impact evaluation. Niger discussed the case of PRODEX (*projet développement des produits agro-pastorales*) and needed to report back on the progress made against the formal engagements during the first Addis conference. The only problem with such training is the general nature of the information provided and the lack of trainers' facilitation techniques.

#### Capacity Building Providers in Niger II: National Training Institutions

Despite the proliferation of local capacity building providers (LCBP) and the provisions of the 2008 Accra Agenda for Action calling for the use of indigenous rather than

external technical assistance, the demand for evaluation services in Niger did not rise as expected. While one may argue that the reliance on donor-funded capacity development was due to the lack of individuals with the right set of evaluation skills and institutions with recognized evaluation expertise in a variety of sectors the country, one may also argue that it was the very aid sector the one discouraging and “crowding out” LCBP. This phenomenon was partly explained by the education policies supported by the World Bank and some other multilateral institutions prioritizing investments in primary and secondary education over tertiary (university-level), as stated by Damtew Teferra in a recent publication discussing the challenges faced by African universities over the last three decades<sup>128</sup>.

Despite the vibrant ReNSE membership, the number of qualified in-country evaluators (national practitioners combining several years of evaluation experience and a good understanding of the major evaluation theories and approaches) was still quite limited at the time this case study was developed. Therefore, as emerged in the course of the interviews with a variety of ECB and ECD stakeholders, the provision of formalized training in evaluation in-country was particularly critical.

The only institution in Niger authorized to train professionals and young graduates in evaluation was The Higher Institute of Evaluation and Perspectives (ISEP). Created in 2006 by a former university professor of economics with fifteen years of work experience as an economist at UNDP and USAID, the master’s degree program offered by the Institute enrolled a total of 120 students over a period of five years, but only a small percentage of

---

<sup>128</sup> Cognizant of the limitations of the education policies supported for in the 1980s and the 1990s, multilateral institutions (e.g., the World Bank) have revised their position on the issue and are currently trying to sustain the development of high-quality programs in universities across Africa (as India has been able to do over the last four decades), as attested by the book “Accelerating catch-up: Tertiary Education for Growth in Sub-Saharan Africa (2009) mentioned by Teferra in his article.

them (12%) completed the program.<sup>129</sup> Regarded as one of the first two evaluation trainings offered in the region, the academic program in question was established at two different levels:

- Master (level 1) for people with a B.S. in sociology, economics, agronomics, and science; and
- Master (level 2) for people who had a more advanced academic level in a variety of domains, including economics, sociology, social science and engineering

The course originated from the realization by its founder (who also happened to be a ReNSE founding member) that a large number of development projects were failing and that therefore it was necessary not only to develop local evaluation capacity through experiential learning programs but also to foster evidence-informed decision-making in the public sector.

During a first phase, the program aimed at creating a favorable environment for the promotion of evaluation. During a second phase, the program started focusing on enhancing local evaluators' skills to master methods and tools. The program (unlike PALAMA in South Africa) was never asked to train public functionaries.

The rationales for establishing such a school were varied. As attested by the School Director as well as other ReNSE members: “The donors’ supply of evaluation capacity building was not sufficient:. It provided no key basic concepts of evaluation and not much guidance on how to conduct data analysis.” Likewise, “the connection between evaluation

---

<sup>129</sup> The total number of students in 2005-2007 was 21 (five pursuing the Master’s at the primary level and 16 pursuing the Master’s at the secondary level). The total number of students in 2011-2012 was 23 (six pursuing the master at the primary level and 17 pursuing it at the secondary level).

and decision-making was not always connected and the fact each donor had its own evaluation policy did not help much.”

Based on the experience accumulated over the last few years, two areas of knowledge in which capacity development seemed particularly relevant for the future were: strategy development and long-term analysis. However, without a national training framework, it was difficult to establish not only how to assess capacity needs but also how to address the identified weaknesses. Experience proved that evaluation capacity needs assessments were of utmost importance. As mentioned by a national evaluation capacity provider:

Students do not know about objectives, logic model, the planning stage. We need to introduce that. The whole discourse (planning and implementation) is missing. If you do not meet your objectives, it is not a tragedy. You need to learn from your mistakes. You need to develop risk analysis and risk management.

#### Curriculum

The school curriculum was specifically focused on teaching so-called evaluation logic, thus clarifying the factors that facilitate evaluative judgment. First came the analysis of the theory of change and implementation strategy (“*démarche stratégique et conception*”). As a result, students learned about project conceptualization and strategy development along with management (“*gestion*”). In promoting the learning of management principles, the curriculum stressed that Results-Based Management (RBM) was not an evaluation, as understood and disseminated across many government agencies. RBM is rather a management technique/philosophy.

Only fifteen students completed the course work and fulfilled their thesis requirements. Among those who graduated, some went to work for an NGO (CARE) and a few others started working for the Government (no specific data on alumni placement is

available, though). Why such a low number of graduates? A few challenges that explain this phenomenon included:

- The curriculum (developed with the help of professors from the University of Senegal, one economist with expertise on Niger, and a professor from Togo) was regarded as too quantitative (econometrics, mathematics, statistics). Student expectations were not very clear - students thought the course would be too qualitative (some just disappeared);
- There were people who just took the minimum course load (Table 12) and did not want to go further with their studies;
- Both professionals and students did not have time to attend on a regular basis: after eight hours of work, they had to stay in school for six additional hours of work per day (eight hours in the past: from 9:00 am to 1:00 pm and then from 3:00 pm to 7:00 pm)

Table 12

*The ISPE Curriculum: Course and Corresponding Credits*

Evaluation Methods (6)	International economics (2)	Development economics (2)
Management (4)	French (2)	Niger economy (2)
Data Analysis (3)	National Budgeting (2)	Forecasting and Analysis(2)
Math (3)	Growth (2)	Arabic (2)
Econometrics (3)	Strategy (2)	IT(2)
Total: 40 credit hours of class + 20 credit hours of Thesis		

## Plans for the Future

Some topics recently added to the new curriculum included:

- International economics (methods effects – effects on prices);
- National budgeting features (it is important that students learn about this);
- Correlation (econometric)
- The use of statistical software packages, such as SPSS

The overall cost was 5,000 francs CFA (US\$ 11.00) for registration and 800,000 francs CFA (US\$1600) for tuition. Fees for the purchase of textbooks were quite low as the School has a virtual library providing students with access to free textbooks. That notwithstanding, the graduation rate had not met initial expectations: as of April 2012, only fifteen students had concluded their coursework and completed their thesis. Among the reasons behind the low graduation rate were: (i) the scheduling of classes was often incompatible with students availability (eight hours of courses during the day until last year when classes started being offered for six hours a day after work; (ii) the lack of professors' punctuality, often due to their busy schedule, with the inevitable effect of disrupting students' learning and creating unnecessary frustration among those attending the courses. The curriculum also appeared to be very much focused on quantitative analysis (thus often not meeting students' expectations) and not always in touch with the current trends; and (iii) the use of old materials: the models discussed in the Niger Economy class, according to some respondents, were the same ones from 1976. The dated curriculum content, though, does not hinder the debates between the professor and the students in class.

Classes normally were held between 8:00a.m. and noon and would then continue from 3:00p.m. to 7:00p.m. Classes would normally last two hours each and some professors (the majority of whom teach at the local university) would cover more classes during the

week (the school director, for instance, taught “Introduction to Strategy” and “Evaluation Methods”). Among the subjects included in the curriculum are communication, Introduction to prospective (e.g., long-term analysis, prospective theory/prospective, economic analysis), accounting, and econometrics.

#### Other Types of Evaluation Trainings Currently Being Developed at the Local University

In 2005, the University of Niamey started going through a new revitalization phase (reform). Instead of offering general Ph.D. or generic courses, the university started developing a new curriculum that would contribute to Niger’s development. One of the new courses that received the most support from different members of the faculty (especially those with extensive M&E consulting experience) at the School of Humanities and Social Sciences was M&E. The need for a “functional” degree was dictated by the relatively low capacity in M&E among NGOs and seemed to fill the gap in technical human resources in a country that has adopted the “project” approach in rural areas as part of its national development policy since 1973. The envisaged master’s degree was inter-disciplinary in scope (the other academic entities participating in the project were the Department of Economics as well as the Department of Geography and Agronomics). All three sponsoring departments offered a generic course in M&E and often exposed students to the topic within the scope of the course on research methodology.

A two-year Master’s degree program is currently under development and expected to start in 2013:

- The first year would be more general and cover fundamental topics such as Development Theory and Project Development;

- The second year would be more action-oriented and would address M&E tools and methodologies in more detail.

At the time the case study was developed, there was some general interest, more so from GIZ (whose office in Niamey often provides local graduate students with internship opportunities) and the Embassy of Spain that supported two other specialty programs in gender and planning. The challenges in setting up this Master's Degree in Evaluation included: a) the lack of qualified teachers (human resources); b) the involvement of three deans and three schools: the School of Social Science, the School of Law and Economics, and the School of Agronomics. (interdisciplinary). That notwithstanding, the success of the program was thought quite plausible, also based on the finding of a recent study on youth unemployment ("*Association nigériennes pour la promotion de l'emploi*") according to which the four most sought-after skills and background on the job market were as follows:

- Sociology
- Agronomics
- Economics
- Geography; and
- The most useful classes were:
  - Statistics
  - Project Development
  - Monitoring
  - Evaluation (different types)



### Capacity Building Providers III: Research and Planning Divisions within Ministries

All the ministries had a DEP (*Direction d'Etudes et Planification*). Each DEP normally had to:

- report to the General Secretariat (a sort of assistant to the Minister) and
- organize (and provide oversight of) all the short- and long-term studies as well as the evaluations commissioned by the Ministry (these may include *prospectives*, that is, long-term programmatic projections/forecasting in line with the government's new long-term vision (replacing the traditional 5-year development plan) PICAC (2 priority year 2011-2012 – *Programme de developpement economic and social*

#### Relevant Contextual Features Contributing to the Development of a Strong Evaluative Culture/ECD in Niger

### The Establishment of Participatory Planning Processes

Despite the lack of trust often characterizing the interactions between government officials and CSOs (see below), Niger put in place mechanisms in 2012 that promised to facilitate the strengthening of links between stakeholders traditionally involved in national planning processes and CSOs, either individual ones or as consortia. As stated by one government representative interviewed in Niamey:

As of 2013, we will need to comply with a new a bill making mandatory for us in the government to have open debates on the national budget with civil society before the bill is discussed in parliament. We also have a specific line in our budget specifically dedicated to the organization of collaborative efforts with citizens' groups. Obviously, we are open to get civil society more involved in the future. However, given their large number, we will not be able to work with all of them. For example, as CBOs are often organized in consortia, I envisage that it will be more practical for us to interact with such collective entities in the future

Similarly, it was recommended by one respondent that a plurality of ministries be involved in the national evaluation discourse: “It is important that each department within a ministry be represented at an evaluation retreat and that they dispose all of an Action Plan, better if accompanied by a multi-year strategy.”

#### Critical Function of the Accountability and Dissemination Principles

Neither the exposure of public programs to external scrutiny nor the review of government officers’ performance seemed to be a common practice in Niger. Therefore, as attested by several ReNSE members at the end of a presentation on institutional performance evaluation:

Civil society is likely to play a critical role in enhancing public accountability and dissemination of information: something that institutions certainly do not like but which they will need to get used to pretty soon, given the increasing donors’ pressures for development effectiveness and value for money.

In particular, it appeared that the practice of disseminating evaluation reports’ findings was not homogenous among in-country evaluation stakeholders. As stated by one independent evaluation practitioner based in Niamey:

The findings of UN-funded evaluations are also disseminated in quite a timely manner. To the contrary, findings of evaluations commissioned by the national governments do not get shared as often as they should and the modalities to recruit consultants that will evaluate public programs are not always clear.

#### Availability of Inspiring Evaluation Champions

Far from being a simple government strategy aimed at promoting better evaluation practices within a given project implementation unit inside a Ministry, the identification of champions was confirmed to have a critical influence on the promotion of an evaluative culture at the national level. In particular what seemed to emerge from the interviews held

with several evaluation professionals in Niger was that the identification of a critical mass of champions with two main characteristics was needed. First, it was necessary for champions to either have an already established affiliation with a certain organization or be likely to develop one (better to combine a cadre of professionals with those project-based officials that had been traditionally most targeted by ECB programs in the past). Second, it was important that champions be identified at several levels within the organization (a single champion was not considered sufficient to promote an evaluative culture within an organization) and within a large variety of organizations, including academia and consulting firms. When asked how they got interested in evaluation, most of the professionals actively involved in the national evaluation discourse mentioned that their first bosses or university professors had been the ones who had exposed them to evaluation for the first time. Some other respondents mentioned that what really got their interest in the evaluation was the political function and values associated with it. As stated by a young professional who had served as a volunteer at ReNSE:

When I first heard of the word evaluation in one of my classes, I got suddenly fascinated by the progressive nature of the undertakings and purposes associated with it: I support social change, transparency and good governance- all of these are values that evaluation allows to push.

#### The Push Towards Professionalization of the Evaluation Profession

As ReNSE presented itself as an association of evaluation professionals, then the Network seemed particularly well placed to work on the definition of evaluation professional standards that all evaluators working in the country would need to abide by. Clearly, this effort would need to build upon the existing professional standards and competencies developed by other evaluation associations, such as AfrEA and IDEAS. As stated by one particularly active national evaluation professional in Niamey:

You need to be consistent with the preaching. There is evolution in the field: new approaches are being developed and the scrutiny from donors is going to be much stronger in the future. It is not possible to improvise evaluation work, what we call “bricolage” or papier-mâché in the French language. You need real experts and you can’t be fearful that you will not meet the requirements. If you possess a real expertise, there should not be any reason to feel that you are going to lose your jobs. This is also an ethical question: A lot of money goes out and (20% of resources go to cover overhead and only 13% of program budgets go to beneficiaries) improving the national evaluation practices will be beneficial to turn the tide and promote further development.

#### VOPEs as a Medium of Citizen’s Control

Could VOPEs serve as citizens control mechanism vis-à-vis public policies? Could VOPEs play the role of what the French call *contre-pouvoir*? As demonstrated by the ReNSE experience, an evaluation network could certainly foster demand for government’s accountability by facilitating a national dialogue on how to (i) measure the performance of public institutions and (ii) use the available evidence to strengthen the coverage, effectiveness and sustainability of public policies. In doing that, though- it was the most respondents’ opinion—it would be appropriate for a network to take a firm stand against any given government or institution. As stated by the ReNSE Coordinator:

Far from becoming a politicized entity associated with a specific political party or ideology, a national network should then make the effort to remain independent and inclusive. An evaluation network should reach out to as many stakeholders as possible (e.g., members of parliament, government officials, civil society organizations, auditor’s general office staff) and promote their understanding and use of evaluation. We need to stay away from conflict and remain *super partes*.

#### Communication, Communication, Communication

As of 2012, ReNSE was still behind in the implementation of an inclusive communication strategy that could focus on the need for both accountability and transparency of information surrounding public programs’ and CSOs’ programs

effectiveness. Two factors seemed to undermine the current efforts, though. First, the large number of representatives of national institutions (e.g., specialists working in the ministries' research and planning units) within its membership made it hard for turning ReNSE into a tool of citizens' control<sup>130</sup>. Second, despite its enhanced visibility, ReNSE did not appear as of yet as a credible broker due to its limited lobbying and communication capacity. That appeared to become possible only when the VOPE would no longer be primarily concerned with its day-to-day tasks and responsibilities (organizational survival management).

#### Recognition of the Evaluation Culture “Intangibles”

The policy-makers interviewed in Niamey felt the pressure of having to deliver results in compliance with the donors' and general population's expectations. The Planning and Research Managing Units existing within Ministries (*Direction d'études et programmation* or DEP) were the governmental actors, that more than others were, tried to follow up on such perceived need for more evaluative data for decision-making purposes. However, despite their good intentions and very high caliber of the people working in them, the DEP did not conduct as many evaluations as they were expected to and focused rather on monitoring. In addition, as such units seemed to play an internal control function, they seemed to be biased in their assessment of public program's performance, thus calling into question the whole utility of their work. Given the existing scenario, it appeared as though a more coherent promotion of governance (e.g., through the use of peer review and external consultants) would need to be conducted within DEP and their respective ministries in the future. As mentioned by one of the DEP coordinators at the Ministry of Finance in Niger:

---

<sup>130</sup> South Africa and Benin have set up M&E systems with a strong citizens' control component (SILPs in Benin and Citizens Delivery Monitoring in South Africa.

People talk about transparency and good governance, but what is needed the most is more information and understanding of data collection and analysis techniques. Should donors be interested in helping Niger further, their efforts should be on developing a culture of evaluation. It is a *behavior*, a *mindset*, not a set of specific skills and bit of knowledge. This kind of support is of utmost importance as governments need to have a culture of evaluation developed within them.

#### Availability of Incentives

Unlike what had happened in Rwanda (performance-based budgeting)<sup>131</sup>, no special incentives (either monetary or non-monetary) were available to government officials in Niger who successfully practiced evaluation and used evaluation findings to guide their decision-making. According to several respondents, it was necessary to make incentives increasingly available in Niger through the systematic inclusion of special budget lines for the conduct and use of evaluation in public programs. In other words, the provision of training was not regarded by some national evaluation practitioners as a sufficiently good strategy to foster demand for evaluation in Niger, as brilliantly explained by one the most prolific ECB scholars:

A lesson without a carrot does not work. For incentives to work, one must ensure that resources are associated with a basic and powerful force: self-interest. (Toulemonde, 1997, p. 24).

---

<sup>131</sup> Based on their satisfactory performance, public servants in Rwanda could receive a bonus, especially if they were able to successfully contribute to the attainment of their ministry's objectives.

## Factors Discouraging the Development of a Strong Evaluativ Culture/ECD in Niger

### Low Level of Government's Trust in Civil Society

National government officials in Niger got increasingly involved in ReNSE activities as well as the broader national discourse on M&E, RBM, MfDR and development effectiveness, since the early 2000s. However, the degree of inclusiveness of the existing M&E processes appeared to be quite limited. One of the reasons explaining the phenomenon, as emerged in the course of data collection, was the lack of trust between the national government and civil society. As stated by one RBM specialist working at the Ministry of Finance in Niamey:

Civil society does not have a distinct evaluation expertise as ReNSE does. When you look at civil society and you consider the possibility of engaging in a discussion with them, you need to distinguish the ones that are good from the one who are not. The majority of civil society organizations do not have any sound technical understanding of evaluation – they are simply looking for funding. Many CSOs are simply opportunistic in whatever they do. In addition, most of such associations are biased: they are politicized and will use their activities and propaganda to attack the government in place. I am of the opinion that you can't really contribute much to the national discourse on evaluation in our country if you do not first get your facts right. What I mean is that you need the knowledge and tools and you need to be ready (CSOs are not always able to do that) to defend and justify your statements on the public sector performance once they release information, more often critical than not of the government, to the public.

How could CSOs get involved in evaluation? Not really sure as I see them as theme-based organizations, that is organizations with a very specific mandate determined by their funders. Yes, I acknowledge the important work that they often do, in terms of service delivery. But, when I think of them, no, I do not really see a soul with a clear ideology and willingness to create and contribute to development.”

The idea that CSOs were not always equipped with the necessary knowledge and experience to participate in a national discourse on how to develop the national capacity further was echoed by several other ReNSE members:

Although civil society defends population rights, there does not seem to be much transparency in the way they get their own funding. Therefore, their independence is not always demonstrated. In addition, they often accuse the people in the government (“*ils chargent les responsables*”) for their alleged wrongdoing and they do not always publish their evaluation. (The government recently tried to conduct a mapping of all the NGOs and CSOs in Niger in late 2010, but a good number of them refused to share their information publicly). I suspect that there is really not a willingness on the part of civil society to promote any form of organizational learning based on the progress or failure of programs (whether public or private).

There is often a sense that evaluation is instrumental to the achievement of political purposes. Evaluations are often conducted by independent evaluation consultants who were not selected based on a transparent competitive process and what I have witnessed in most cases is the lack of a real stakeholder participation.

Far from suggesting that CSOs become servile to the government, what I would like to see is that CSOs, rather than attacking the government, support the different ministries and other branches of the government by both collecting more data on the actual state of public program and services and sharing such information in a timely manner. Instead of protesting, as if they were the messengers of foreign and external views on national politics (as is the case with those CSOs funded by international NGOs), CSOs should work harder to develop their own statistical capacity and improve their own understanding of evaluation methodologies. What is even more important is that they also, based on the resources available to them, work with national institutions to develop their capacity further.

Due to this low level of trust, the CSO representatives interviewed denounced that lack of dissemination of findings related to the evaluation of public government programs:

The Evaluation units within ministries are the black beasts. When evaluators like us call them to have their data, they are always afraid that we are going to use it in order to penalize them and that the data are going to be used to mount negative propaganda against their public programs. The whole principle of program improvement and organizational learning is not understood. Far from being regarded as a pedagogic tool, evaluation is still considered an inspection tool *at all levels*, not only within civil society.

The lack of an effective decentralization of the M&E function



Despite the existence of decentralized entities since 1991, the decentralization process in Niger became effective only in 2005. In addition, although Niger had locally elected officials (*conseillers municipaux*), the powers transferred to the latter were very limited, unlike the South Africa case discussed earlier in this chapter.

#### The Relatively Weak Statistical Capacity

As stated by a former ReNSE coordinator and national evaluation practitioner with some of the country's widest expertise in the use of quantitative methods:

Statisticians do various quantitative analyses but they are not very strong and only a very few of them understand randomization. Recently, the evaluation of the population policy was a story of the policy (purely qualitative narrative) and some great opportunities were missed. For example, a control group could have been used to examine the changes in contraception use at the national level. In addition, agronomists, sociologists and university professors conduct data collections evaluation, but their data analysis skills remain quite weak.

#### AfCoP-ization of the evaluation function

The overlap in membership between ReNSE and the African Community of Practice on Management for Development Results (AfCoP-MfDR) as well as the prominent role played by the Ministry of Planning and Finance in the promotion of M&E in Niger certainly contributed to the integration of result-based practices in Niger's national planning and management processes over the last decade. In particular, the creation of the National Chapter of the AfCoP-MfDR in Niamey, under the auspices of the African Development Bank and the World Bank in 2008, represented a cementing factor in the strengthening of the national development and evaluation community. As attested by a Ministry of Planning official:

RBM is a strategy that measures performance to get to results. For example, a school is not the result of an education project but rather a deliverable that serves higher level goals, such as enhancing students' learning practices and

boost their academic results. RBM pushes people not only to think about what we do but even question at times our own assumptions about how development works. RBM is truly a government's responsibility: it is a tool that all government should use to guide the development engine.

RBM-based culture<sup>132</sup> was so pervasive in Niger that most respondents met in the course of the data collection described evaluation as if it was RBM<sup>133</sup>. In other words, most of the government officials and evaluation practitioners interviewed in Niamey shared the understanding that evaluation data were primarily aimed at measuring the level of progress attained by a project or program against a set of objectives included in a pre-defined logic model. As a result of that, it appeared that the other evaluation principles commonly referred to in international development evaluation (relevance, efficiency and impact—both expected and unexpected) were completely dismissed.

That the majority of the evaluation practitioners in Niger still regarded evaluation as an accountability—and not learning-driven activity—was confirmed by a variety of statements collected in the course of the field data collection in Niamey, such as: “Evaluation is about making sure that the expected objectives have been produced’ or “Evaluation is about verifying compliance between results on the ground and program design.” That notwithstanding, it was important to mention that, if it had not been for RBM, many government officials, especially planning specialists and those working in Ministers’ Research and Planning Units (in French DEP – *Division d’Etudes et Planification*) would have not learned about M&E. Also, if it has not been for RBM, many government officials might have not been able to understand the relevance of integrating M&E it into the formulation

---

<sup>132</sup> Results-based management is a management strategy by which all actors on the ground, contributing directly or indirectly to achieving a set of development results, ensure that their processes, products and services contribute to the achievement of desired results (outputs, outcomes and goals). RBM rests on clearly defined accountability for results and requires monitoring and self-assessment of progress towards results, including reporting on performance. (United Nations Development Group: Results Management Handbook 2010, p.5)

<sup>133</sup> M&E is one of the RBM pillars but the purpose of evaluation cannot be associated only with the principles of accountability and compliance which RBM rests upon.

of their respective action plans as well as need for developing monitoring progress indicators and roadmaps to attain the envisaged results. As reiterated by a few of the government officials who are actively involved in ReNSE:

“Evaluation is instrumental to the management function and evaluation is now part of the project cycle.”

“RBM and Evaluation are not different content-wise. They complete each other, they both pursue the results.”

Based on the previous findings, a question was then raised as to whether RBM did a disservice to the promotion of an evaluative culture. On the one hand, it was true that MfDR served as the point of entry for the dissemination of key development principles included in international covenants (e.g., aid effectiveness in the Paris Declaration, Relevance and Sustainability in the OECD list of evaluation criteria, donors’ harmonization and ownership of development processes in the Accra Agenda for Action). On the other hand, the evaluation logic (that is, the rigorous determination of the merit, worth and significance of any given program and policy against a set of agreed upon principles) did not seem to be well understood and practiced by those very same stakeholders (e.g., chiefs of divisions within ministries or members of parliaments) who were more likely than other national stakeholders to influence the successful promotion of an evaluation culture. The limited use of evaluation in Niger, despite the principle of mainstreaming M&E in the national planning processes, seemed to confirm the distortive effects of the proximity (often so close to be seem an overlap) between RBM and evaluation, which prevented national practitioners from fully appreciating evaluation’s traditional formative and summative functions.

## ReNSE Membership: Do Ownership and Inclusiveness Always Go Together?

The state of evaluation in Niger was quite different than the one in SA for different reasons. On the one hand, Niger was characterized by a relatively lower degree of institutionalization of the evaluation function (South Africa instead had already established a Performance Monitoring and Evaluation Department and the government in Pretoria, under the aegis of the President, conducted annual public assessments of departments' performance in compliance with the provisions of a National Evaluation Strategy). On the other hand, the proximity between ReNSE and the country's institutional actors appeared much greater than the one between SAMEA and the DPME in South Africa. However, the relationship between the Network and the national government had not always been characterized in such positive terms.

During the ReNSE informal conceptualization and development phase (late 1990s), the national ownership of the evaluation function in Niger was particularly low, due to the heavy involvement of UNICEF and UNDP in the set-up, coordination, management and financial support of the evaluation discourse. It was only in 2007 and after the formalization of the Network in 2009 that the national ownership of the evaluation function seemed to have grown, as attested by the larger share of government officials within the ReNSE membership and the first attempt to draft a national evaluation strategy)<sup>134</sup>. Interestingly, the degree of ownership over the evaluation function seems to be inversely correlated to the degree of inclusiveness of the national consultation processes.

---

<sup>134</sup> As a result, not only does ReNSE promote the practice and use of evaluation mainly within the national government but it also contributed quite actively to the drafting of the National Evaluation Policy as well as to the evaluation of the National Poverty Reduction Strategy.

## Pros and Cons of Formalizing an Informal Network: The ReNSE Experience

In 2007 ReNSE leadership, with the support of several development partners, decided to acquire a juridical status. The formalization process, whose completion took almost two years (ReNSE became formalized in 2009), spurred some debates among ReNSE members on both the risks and advantages associated with formalization. Learning more about the exchanges that occurred within ReNSE membership during the two years of transition (from an informal network to a formalized entity) was quite useful as it appeared to provide some interesting ideas that would be able to (i) inform the planning of other VOPEs considering whether or not to formalize, and (ii) provide some tips on how to mitigate the risks associated with formalization and to manage effectively the critical organizational development changes spurred by it (Guijt, 2008). Below is an overview of all the pros and cons of formalizing a national evaluation network.

### Arguments in Favor of ReNSE Formalization

#### *Financial Benefit*

At a time when ReNSE development partners were planning a gradual phase out from the Network's programming, the Network's financial and institutional sustainability seemed to be threatened for the very first time. Therefore, formalization was rapidly identified as the most obvious strategy to pursue: it would make the network eligible to apply for grants (e.g., ReNSE coordinating committee staff would acquire the right of signature) and keep "playing according to the rules of the donor-grantee development game" (Yachkaschi, 2010).

### *Legitimacy to Provide Evaluation Services*

Strictly related to the prior argument, this one was in favor of the network's acquisition of a network's juridical status, as it would enable ReNSE to provide (and not only receive as had been the case in the past) professional evaluation services;

### *Increased National Ownership and Professionalization of the Network leadership*

Due to the "zero budget" principle (according to which no overhead costs could be covered by external funding source and individuals involved in the network coordination were to do so on a purely voluntary basis), many of the ReNSE members believed that the formalization would pave the way for more national professionals (rather than UN officers) more qualified and adequately remunerated, to lead the Network.

### *Enhanced National and International Credibility*

Some other members were of the opinion that a network with enhanced planning and financial autonomy (e.g., including the ability to both apply for external funding and allocate resources for priorities identified by its own network membership) would benefit from increased visibility and recognition in international settings;

### *More Efficient Administrative Processes*

As formalization is always accompanied by legalization, many members argued that an official registration of the Network's juridical status would contribute to the strengthening of ReNSE internal governance structures. As a result of formalization, ReNSE members believed that they would greatly benefit from more orderly and transparent internal processes, such as those more closely related to decision-making and

dispute settlement. At the same time, the set-up and dissemination of official (agreed in writing) procedures was seen as the basis for a more diligent behavior among members.

#### Arguments against ReNSE Formalization

##### *Time Consuming Process*

The discussions over the formalization of ReNSE were very time-consuming and seemed to have paralyzed what was by then quite a vibrant VOPE. In the ReNSE case, concurrently with the discussions held in 2007-2009, the exchanges on the website as well as the work of the thematic groups were temporarily interrupted. As stated by a former ReNSE coordinator, it was necessary that similar initiatives in the future would involve a more restrained group of members:

If I had to do it again, I would keep the membership's involvement in the conceptualization and implementation of the formalization process to a normal acceptable level without making it the focus of our work for so many months. In order to do that, I would delegate the tasks to work on the formalization to a much smaller group of members.

##### *Membership Decrease*

The institutionalization of membership fees discouraged many individuals from remaining involved in ReNSE activities. In Niger, membership dropped from 300 to 38 (this number was calculated based on the number members with voting rights participating in the General Assembly in April 2012)<sup>135</sup>. In order to respond to such drastic drop in membership, the ReNSE coordinating committee in the course of the last General Assembly, implemented the decision, agreed upon by all the attending members, that those non-paying members would still be able to participate in official meetings;

---

<sup>135</sup> Some might argue that in an informal network, it is always a very small number of the existing that actively participate in all activities. In the case of formalization, paying members are no other than those very same truly engaged members.

### *Slow-down in the Pace of Activities*

Following the formalization, ReNSE activities seemed to take longer than usual to be implemented. As stated by one of the ReNSE members:

While formalization is adequate if you are planning to set up a research center or a consulting firm specializing in evaluation, creating borders (e.g., providing who qualifies to be a member and how much membership fees are), it also prevents people with fewer economic resources and a more modest appreciation for structured organization from either actively volunteering in ReNSE activities or contributing to the overall internal debates.

### *Fragmentation within Membership*

The distinction among founding, active and partner members (article 8 of the ReNSE constitution) was identified as the possible cause for fragmentation within the VOPE's membership. Therefore, it was recommended that the Network provide an open platform for all in-country practitioners with a special interest in evaluation regardless of their level of seniority or their ability to pay the membership fees.

### *Higher Operational Costs*

Once ReNSE was formalized, office rent needed to be paid. In addition, it was argued that the coordinator functions, which had never been remunerated before, should be recognized with some form of monetary compensation.

### *Reduction in the Involvement of Individual Officers Working for International Development Partners*

As the standard rule was that donors could not be members of formalized national networks, UNICEF individual staff could no longer be registered as individual members but they could only join the Network as an institutional entity (as per Article 4.2 of the ReNSE Statute).



### *Possible Lower Level of Donor's Harmonization in ECD*

So long as ReNSE remained informal, its meetings took place in different venues each time, thus enhancing shared ownership of the related planning and coordination processes among the several development partners supporting ReNSE. However, as the habit of holding ReNSE meetings at the offices of development partners was interrupted, the risk emerged that donors would be less involved both individually and collectively in ReNSE activities.

### *The Rigidity of Membership Categories Hinders Creativity and Inclusiveness*

Some of the ReNSE members, especially those who did not pay the membership fees, were vocal about the principle that a network, by virtue of its informal nature, should not be formalized. They also added:

In an effort to promote an inclusive dialogue, it is important that membership be sufficiently permeable so as to include different membership categories. However, formalization does not really help to do that. In the ReNSE case, a good model that could have been replicated was that of the CoP-MfDR, characterized by a lighter structure and yet quite selective criteria to join. Its coordination committee included 10 individuals (one coordinator and one rapporteur). Membership was not mandatory and all its undertakings were pursued in a very participatory manner.

### *The Future of ECD in Niger*

Based on the interviews held in Niamey with a variety of evaluation practitioners both in the public and private sectors, the main features of desirable capacity development activities to implement in the future were identified.

First, it was stressed that ECD in Niger would certainly include workshops of variable duration (1, 2 or 3 days), but would not be the most desirable or most central activity to be implemented. In particular, besides questioning the effectiveness of

workshops in more general terms, a few respondents lamented that a) such short-term training did not always address specific capacity needs and b) the per diem associated with the participation in such training was often the only motivational factor behind attendance. Second, a renewed focus on university programs appeared to be one of the most favorite topics among respondents. In particular, the strengthening of existing university programs in evaluation as well as the development of new evaluation courses (three or six months in duration and held over the weekend) both for evaluation practitioners and commissioners was recommended. According to one of the ReNSE members:

A new university evaluation course should be created: delivered by a well-established evaluator, this academic activity should entail homework and the wide use of case studies. In addition, participants should be given the opportunity to receive a certificate upon the successful completion of all the modules. The only issue associated with this recommended option is that, were an international expert to offer the course, she/he would not be too familiar with the local context. Two different specialty courses should be made available: a shorter course for evaluation managers/commissioners and a longer one aimed at evaluation specialists.

Third, several respondents suggested that a national event aimed at sensitizing policy makers further on the role of evaluation, and led by a group of MPs (not always the same group of evaluation supporters within the General Directorate for Evaluation) would need to be organized. To this end, the evaluation of public policies and programs, as well as the evaluation of the public administration's institutional performance, were regarded as particularly useful topics for a conference. Fourth, respondents suggested that future evaluation programs should focus on the fundamentals: a) data analysis; b) evaluation methods; c) the use of specialized software; d) sampling; e) the monitoring of the so-called "Black Box"; f) project development; g) use of evaluation of accountability purposes; h) the difference between evaluation and auditing; i) the development of evaluation ToRs; and l)

meta evaluation; m) evaluation as opposed to feasibility studies or ex-ante impact assessments; and n) ethical standards in evaluation. Fifth, respondents suggested that future M&E Capacity Development should increasingly include coaching and mentoring activities. It was suggested, for instance, that those ReNSE members with greater technical expertise could provide a 3-day introduction on a specific evaluation topics to the rest of the members and then provide monthly refresher courses for a trial period of one year and that the ministries, if satisfied with the results, could provide more resources for similar activities within several ministries in the future. One more process-related suggestion was that ReNSE paying members be given priority to take part in these formal trainings (at the time of data collection this was already the case within the scope of training organized at the DEP level) and that special attendance quota be established in favor of members from the private sector and civil society. Sixth, it was suggested that the ReNSE Thematic Groups be given the opportunity to follow evaluation training in their specific sector of interest and that, in the absence of in house expertise, professionals be invited to present for one or two hours on a given topic chosen by a scientific committee. Seventh, it was suggested that some cost-effective ECD activities in the future be implemented a trial basis. These included: a) the involvement of students who had received the national Master's degree in evaluation; b) the involvement of professionals providing their services on a *pro bono* basis in exchange for public recognition; c) the organization of information campaigns on accountability and evaluation among high school students; and d) the provision of support by ReNSE to agencies (e.g., Transparency International or ADDH) conducting crucial data collection nationwide.

## CHAPTER 7

### CASE STUDY 3: THE DEMOCRATIC REPUBLIC OF CONGO MONITORING AND EVALUATION ASSOCIATION (ACoSE)

#### Introduction

Due the vast array of development projects implemented in the DRC in a number of areas (from agriculture and education to governance, health and private sector development), the number of evaluations conducted in country (mostly by foreign donors and consulting firms) was quite remarkable for over a decade. However, as most development projects implemented in the DRC qualified as humanitarian assistance or emergency projects<sup>136</sup>, most of their evaluations were principally focused on describing outputs and short-term results, rather than on assessing relevance and impact (Valery & Shakir, 2005). As a result, the demand for evaluation in the DRC was quite limited<sup>137</sup> and so was the supply of evaluation services by both national practitioners and institutions (whether public or private). Furthermore, although monitoring of development projects seemed to be conducted on a more regular basis, the quality of the corresponding data analysis was often questionable. All such weaknesses had two primary repercussions.

---

<sup>136</sup> Most of such interventions were aimed at mitigating the disastrous effects resulting from decades of genocide and internal conflicts, which the DRC, despite having signed the Lusaka Peace Accords on July 7, 1999 was still recovering from by the end of 2012.

<sup>137</sup> One example of the limited role assigned to evaluation by the national government is the lack of a M&E section in the Health Sector Development Plan.

First, and more importantly, the lack a national evaluative culture hindered the development of an accountability culture in the public sector<sup>138</sup> and discouraged attainment of better social and economic conditions<sup>139</sup> among the Congolese population. Second, the absence of a critical mass of evaluation practitioners discouraged the development of a national evaluation community. That notwithstanding, the Congolese government and a variety of development partners over the last decade (especially in response to the Millennium Development Goal Campaign) made some efforts to promote better M&E practices nationwide. The creation of the Congolese Monitoring and Evaluation Association (Association Congolaise de Suivi/Evaluation or ACoSE) in 2007 was just one of the initiatives aimed at altering the status quo and promoting evidence-informed decision-making for better development results.

A case study developed around the experience of the VOPE in the DRC is presented in this section. The case study consists of four main parts. In the first, the four most salient phases of the Association's history (Conceptualization & Formalization, Institutional Engagement, Stagnation and Opportunistic Revitalization) were identified. In the second part, the ACoSE organizational capacity was described, followed by a detailed description of the Association's role in the promotion of a national evaluative culture. In the third part, the contribution of other national ECD stakeholders was presented. In the fourth and last part,

---

<sup>138</sup> The poor state of the Congolese administration's accountability was attested, among the others, by the following challenges: (i) the unknown number of employed public servants; (ii) the large share of the public administration payroll over the GDP (12%) and the public expenditure (nearly 50%); (iii) a large but unknown number of specific bonus and allowance schemes for different civil servants in different ministries; (iv) the lack of a merit based civil service; (v) the large discrepancies between central government and provincial pay, without a clear rationale, and; (vii) an entrenched system of patronage and placement of highly paid political appointees (World Bank, 2010)

<sup>139</sup> One of the most populous countries in Africa (according to the European Commission, the country's population is estimated to be nearly 77 million people), the DRC continues to be ranked at the bottom of both the Human Development Index (UNDP, 2011) and the IFPRI's 2011 Global Hunger Index (IFPRI, 2011). With a per capita GDP of below US\$ 200, DRC remains the country with the highest infant mortality rate even though it has considerable wealth in mineral resources.

an overview of the factors that either enable or hinder the development of an evaluative culture in the DRC evaluation efforts was provided. Through this case study, conducted in May-June 2012, it was possible to identify a few lessons learned that could inform the design, management and evaluation of future ECD programming in challenging socio-economic and political contexts characterized by the presence of a relatively young VOPE.

### ACoSE History

Based on the reconstruction of the most salient events of ACoSE's history since the inception of its activities, four main phases were identified.

#### First Phase: Conceptualization & Formalization (mid-2007-mid-2008)

The Congolese Monitoring and Evaluation Association (ACoSE) was founded in May 2007, in the aftermath of the AfrEA conference held in Niger. It was during that same conference that the few Congolese attending the event,<sup>140</sup> supported by UNFPA, started discussing the idea of creating a national evaluation association in the DRC. Since the beginning, ACoSE membership was quite varied. Members were evaluation practitioners from the public and private sector as well as researchers and university staff (as stated in the preface of the ACoSE statute).<sup>141</sup>

#### Second Phase: Institutional Engagement (mid-2008-mid 2010)

Following its registration as a non-profit in mid-2007, ACoSE started working very closely with the Ministry of Planning and even became one of the few Congolese agencies

---

<sup>140</sup> This group included the current ACoSE President (serving as a consultant in the DRSP unit within the Ministry of Planning) as well as the coordinator of the ACoSE Civil Society Working Group, and a UNICEF M&E Officer.

<sup>141</sup> As of May 2012, 50% percent of members worked in government, 40% served as planners and M&E officers in development organizations, while the remaining 10% were either unemployed or independent consultants.

designated to disseminate the National Development Strategy Paper<sup>142</sup>. During this phase, ACOSE developed a special collaborative effort with the DRSP Unit within the Ministry of Planning and it also formalized its relationship with the national CoP-MfDR. Some of ACOSE key achievements during this phase included: a) the creation of a few thematic groups that consisted of members sharing the same thematic and sectorial interests<sup>143</sup>; b) the organization of three workshops<sup>144</sup> by one of the Association's Permanent Committees (the training committee) which mainly catered to the paying members and three young graduates in 2011; c) the creation of a website (www.ACoSE-rdcongo.com) with the idea that all the evaluation reports conducted in the DRC would be posted on it (the website was never really used); and d) the organization of a statistics training in collaboration with the Higher Institute of Statistics (*Institut Supérieur des Statistiques*) in 2011. That notwithstanding, ACoSE remained unknown to the general public. Therefore, in an effort to enhance its visibility and increase its membership, ACoSE reached out to civil society and joined the National Civil Society Platform.<sup>145</sup> In addition, the Association became a signatory of the Congolese Chart of Civil Society and strengthened its ties with both AfrEA and UNICEF<sup>146</sup>.

---

<sup>142</sup> In particular, ACoSE started collaborating with the UPPE (*Unité de Pilotage du Processus de l'élaboration de la Stratégie de Réduction de la Pauvreté*), the governmental entity that was the most heavily involved in the promotion of the evaluation function in the DRC. The role played by the UPPE staff was so critical to the implementation of the PRSP (2012-2017) that UNFPA decided to second a M&E specialist to it, with the specific task of assisting the governmental officers in the development of a data collection plan over the next two years

<sup>143</sup> The thematic groups were complementary to the three Permanent Commissions established by the ACOSE statute: a) the Conference, Seminars and Meetings Committee; b) the Training and Exchange Committee; and c) The Newsletter and Other Publications Committee.

<sup>144</sup> The trainings focused respectively on a) the dissemination of RBM principles and implementation strategies; b) M&E (including main approaches, roles and responsibilities and guiding technical principles); and c) the utilization of Microsoft Project

<sup>145</sup> The Platform consisted of twenty different civil society groups (each of which corresponded to either a specific sector -such as agriculture or health- or a cross-cutting thematic area – such as governance, civil society and development effectiveness).

<sup>146</sup> UNICEF had provided a computer, a printer, a photocopy machine and a projector to the ACoSE office right after the formalization.

### Third Phase: Stagnation (2010-2012)

As of mid-2011, ACoSE started experiencing a series of organizational development challenges. First, despite the availability of two Action Plans (one for the period of 2009-2011 and one for 2012<sup>147</sup>), the Association could not implement a large number of planned activities due to the lack of financial resources. Second, day-to-day operational decisions were always made by two or three members without any type of oversight from the General Assembly and members started complaining about the legitimacy of internal processes. In addition, the preparation of the Associations' annual reports took much longer than expected and the synthesis of the three workshops held in 2011, despite the final reports having been made available by the trainers in a timely manner, were not finalized as of June 2012. Third, although ACoSE had recruited a communication specialist (in charge of publishing an information bulletin), no communication strategy was in place and no website, despite the initial plan, was created. Fourth, although the ACoSE President had been hired by UNFPA to be seconded to the PRSP unit within the Ministry of Planning, ACoSE at the organizational level had not promoted any specific activity aligned with the PRSP. In addition, no specific knowledge sharing on this topic has taken place among the Association's members. As reported by one of the ACoSE members:

The association does not own its mandate and has failed to organize any event (e.g., a press conference) on MDG." Fifth, the absence of General Assembly meetings in 2012 severely compromised the fulfillment of the Association's mission. Sixth, ACoSE seemed to have a developed governance structure, including a General Assembly as well as a Managing Committee (Comité Directeur). However, despite the formal provisions included in the Association's statute, the General Assembly only met a couple of times over the last three years. In reality, it was the Managing Committee (not the General Assembly) that carried out the Association's day-to-day tasks and made the decisions on behalf

---

<sup>147</sup> None of them was available at the time the interviews with the associations' president and other members took place in Kinshasa.



of the whole membership. Furthermore, the Managing Committee was in place for over five years instead of the three envisaged in the association's statute.

#### Fourth Phase: Opportunistic Revitalization (July 2012-present)

Despite the challenging political and social environment, ACoSE started planning a series of initiatives that would take advantage of some new opportunities to become available after the political elections organized in late 2011. Some of the planned activities to be implemented during this fourth phase included:

- The involvement of some of the newly elected MPs in evaluation-focused events;
- The organization of awareness raising events (first phase of a larger sensitization program aimed at national policy-makers);
- The development of a data collection and technical capacity building strategy within the scope of the PRSP national planning and evaluation processes.

In the pursuit of such actions, it was apparent that ACoSE would be more successful by aligning the Association's activities with five key objectives included in new Government Programme (2012-2016): job creation, infrastructure, health and education, water and electricity, and housing.<sup>148</sup> Likewise, the government's call for a national anti-corruption campaign seemed to provide ACoSE the opportunity for becoming a national champion of M&E for both accountability and learning purposes.

---

<sup>148</sup>The Congolese Prime Minister argued that such program, known under the name of "Le Programme des Cinq Chantiers" (The "5 Construction Sites" Development Program), would help place Congo right next to a variety of other middle-income countries. In particular, the forecasts were that the DRC would attain a gross national income of between 976 and 11,906 U.S. dollars (like South Africa Brazil and Argentina) by the end of 2016. Source: [www.lobservateur.cd](http://www.lobservateur.cd) May 8, 2012 (last accessed August 10, 2012).

## ACoSE Capability Assessment

The findings of the ACoSE capability assessment, based on the measurement of five different types of organizational capability (Capability to commit and act, Capability to generate development results, Capability to relate, Capability to adapt, Capability to integrate), are presented in this section.

### Capability to Commit and Act

Overall, ACoSE demonstrated an average level of capability to take action. It definitely had a well-articulated mission and an exhaustive list of objectives to achieve (Box 4). However, most of the activities included in the two Action Plans (2010-2011 and 2012) could not be implemented due to lack of budget resources.

Likewise, although the Statute provided for General Assembly meetings to take place at least twice during a year (art. 13), no meeting was organized for a period of at least two years. As a result, the Board of Directors became the only effective deliberating organ within the Association. That had the inevitable effect of overextending the Board's mandate (from 3 to 5 years) in non-compliance with the Statute provisions.

Despite the challenges experienced at the organizational development level and notwithstanding the failed implementation of the Action Plans, ACoSE demonstrated a decent ability to mobilize resources. Although insufficient to cover the implementation costs associated with the two ambitious Action plans, the Association was able to enforce the payment of membership fees (both the \$50 initiation fee and the \$10 quarterly fee) and received some in-kind support from UNICEF (computer, photocopy machine, printer, projector) and UNFPA. As was the case for ReNSE in Niger, ACoSE benefited from the active involvement of a full-time volunteer. A young economist by background, this

**Box 4****ACoSE Objectives**

- Promote the culture of M&E in the DRC among the general population and decision-makers;
- Set up a platform for meeting, interaction, and exchange of M&E information and experience;
- Promote the improvement of evaluation quality and dissemination;
- Improve the quality of M&E in the DRC and turn it into a reference both at the national and international levels;
- Contribute to national development through the adoption of new M&E approaches;
- Contribute to advancing M&E techniques and methods and promote adherence to ethical and procedural rules enhancing evaluation quality and use;
- Promote African evaluation standards;
- Strengthen its members' capacity;
- Contribute to the adoption of a results and accountability culture;
- Contribute to the development of evaluation knowledge, skills and methods;
- Promote the transfer of skills and support the development of trainings provided by academic and professional institutions;
- Develop exchanges among Congolese experts, and between them and their foreign colleagues;
- Promote the dissemination of information and any other work concerning evaluations as well as the institutions and actors working in this area;
- Create a database of national evaluators that may be selected to carry out monitoring and evaluation assignments at national and international levels;
- Promote the capitalization of knowledge accumulated through the evaluations;
- Facilitate the recognition of specific evaluation skills;
- Organize thematic working groups at the provincial, national or regional levels
- Organize conferences, seminars or workshops in partnership with public or private structures;
- Establish links with other similar associations

Source: ACoSE Statute (2007)

individual was responsible for receiving members visiting the office and interested in consulting either the literature available at the office or the online resources accessible from one of the four computer stations installed in the Kinshasa office. Unfortunately, the

Scientific Committee established within the Association never became operative. Its purpose was to ensure the Association's compliance with the international M&E standards.

#### Capability to Generate Development Results

Given the limited financial resources, ACoSE could not generate as many results as originally envisaged. The three main uses of the Association's resources (the purchase of office furniture, the set-up of a website that never really became operative and the organization of three workshops in 3 years) was an indication of the limitations not so of the Association's fund-raising capability but rather of its planning and implementation skills. The young volunteer assisting with the ACoSE day-to-day operations could have been more instrumental to the Association if some more responsibilities had been assigned to him, such as the task of serving as the ACoSE spokesperson in various meetings with government officials, development partners and civil society. Likewise, it was suggested that this individual revitalize the website and disseminate vacancies information among the Association's members (as of May 2012, membership include 32 paying and 20 non-paying).<sup>149</sup>

That the Association could not generate development results was also confirmed by the lack of any concrete achievement resulting from its participation in the Civil Society Platform. That notwithstanding, several ACoSE members, because of their professional affiliation with several key national ministries, actively participated in a number of meetings organized as part of the national PRSP planning process<sup>150</sup>.

---

<sup>149</sup> According to the African Evaluation Association website, the number of members rose to 42 (<http://www.afrea.org/?q=user/26> ; last accessed on October 17, 2012)

<sup>150</sup> ACoSE members' involvement in the process seemed to be more effective than the contributions provided by CSO in the mid-2000s when the model of interaction between government and civil society was centered around consultation more than participation (Morazán, P., & Knoke, 2005; Ntalaja, 2005)

### Capability to Relate

ACOSE was one of the organizations signatories of the Congolese Chart for Civil Society. Likewise, it established close ties with the UPPE unit within the Ministry of Planning and a few of the national universities. The Association also had pretty good ties with UNICEF, UNFPA, AfrEA and the French Evaluation Society. The variety of the ACoSE memberships (50% government officers; 40% implementing agencies; 10% unemployed) enhanced the Association's capability to relate.

### Capability to Adapt

ACoSE demonstrate a discreet capability to adapt over the last few years. However, the Association's plan to explore collaboration opportunities with newly elected MP demonstrated a certain openness to unforeseen scenario and possibilities.

### Capability to Integrate

Overall, ACoSE programming demonstrated itself to be partially responsive to the vast array of interests and needs emerging among its members since the inception of its activities. The creation of thematic groups within the Association and the establishment of relatively close links between such group and a number of CSO working in a variety of sectors were certainly very relevant to ACoSe's visibility and social legitimacy. However, the short lifespan of such initiatives was an illustration of the Association's relatively low capability to solve tensions and bring members of different backgrounds together on a sustained basis.

## Other ECD Stakeholders in DRC

### World Bank

The US\$30 million 5-year Building Sustainable Capacity (BSC) project funded by a World Bank technical assistance grant, was a good illustration of the donors' renewed involvement in national capacity building efforts<sup>151</sup> in the DRC. This single intervention was an example of the new development partners' strategy implemented in the DRC in the aftermath of the Development Effectiveness Forum held in November 2009 and better known as "Kinshasa's Agenda" (*Agenda de Kinshasa* in French). Endorsed by the Congolese government, this new strategy called for a re-orientation of the donor approach to supporting capacity development in the DRC, that is, the shift from ad hoc interventions through disparate development partner programs towards an integrated approach to supporting local institutions, thus allowing for multiplier effects. In response to the World Bank's increasing pressure to fund capacity building activities (Priority Actions) specifically focused on such principles as Leadership, Public Financial Management and Procurement (e.g., RBM), Organizational Behavior and Ethics during the first three years of implementation<sup>152</sup>; the Congolese government committed to carrying out two main tasks. First, the coordination of development interventions implemented by national entities and externally funded. Second, active participation in the planned evaluation endeavors. This undoubtedly introduced some new dynamic in the country-level ECD programming. As stated by one of the SENAREC officials met in Kinshasa:

---

<sup>151</sup> According to the BSC project document: "Addressing governance concerns and building up capacity to create sustainable institutions, both governmental and non-governmental, is a condition sine-qua-non for the development of the DRC and its survival as a unitary state" (World Bank, 2009).

<sup>152</sup> This first phase was characterized by the use of both coaching for change management (a pool of regional coaches was formed) and the Rapid Results Initiative methodology, that had been successfully piloted by the World Bank in Morocco, Madagascar and Burundi (World Bank, 2010).

There are all sorts of government programs funded by external donors but the government can't evaluate such programs. Why? Because a national government cannot evaluate donors' interventions. As a result, many evaluations are conducted by the donors themselves with the primary objective to obtain validation of the original design and implementation modalities rather than to assess the effective relevance and impact of such interventions.

Interestingly, upon completion of this priority phase, the Project would consist of more programmatic actions based on “annual plans, sound fiduciary management, and effective reporting that would require disbursement to be closely linked to results” (World Bank, 2010, p. 6). The CBS project appeared to be all the more innovative as it first acknowledged the critical role played by CSOs (e.g., religious organization) in the delivery of social services, and then it attempted to revamp the national academic training system not only in Kinshasa but in other areas of the country, such as Lubumbashi and Kisangani<sup>153</sup>.

Based on the findings of a national diagnostic of the Congolese Public Administration conducted in 2009 by a national expert in preparation of a conference on aid effectiveness, this project was quite promising. The most striking finding of this assessment was that the capacity of personnel in a large number of governmental agencies was particularly low and that, in order for effective reforms to take place, a sustained effort was needed as soon as possible to both support and qualify further civil servants in leadership positions. Recognizing the critical role of national leadership in generating a sustained effort to enhance the capacity at all levels within the government, the BSC has as objective to fund a leadership program with selected key ministries. The program was particularly targeted to

---

<sup>153</sup> The Project resources will be allocated to fund the following components: Strengthening leadership and creating basic results-based management capacity (US\$4.30 mil), Strengthening Core Functions of Public Administration (US\$7.10 mil); Support Training Centers to Design and Implement Programs of Excellence (US\$13.30 mil); Strengthening SENAREC and building a quality control system (US\$5.20 mil).

the office of the Prime Minister, key functions in line ministries and some centers of excellence<sup>154</sup>, including the offices of Permanent Secretaries (*Sécrétaires Généraux*).

Similarly, the project also engaged CSOs both in the design of the capacity building programs and the monitoring and evaluation of this project. Here, also, the World Bank Institute WBI was actively involved as well as in establishing South-South networks, Rapid Response, and the Global Development Learning Network (GDLN). In support of this last strategy, the project also provided support to six centers of excellence<sup>155</sup> that would be in charge of the training of stakeholders both in the public and private sectors.

#### Other Entities Outside the Government

##### *UNFPA*

One of the strongest UN supporters of the evaluation function in DRC was UNFPA, which, for years, had made a special effort to enhance the technical evaluation capacity of its staff through the organization of 3-5 day annual workshops. Interestingly, despite UNFPA interest in aligning its capacity building efforts with those of other agencies and development partners, most of the M&E workshops it sponsored were usually geared towards the enhancement of the technical, communication and financial capacity among its decentralized staff as well as the improvement of their understanding of RBM practices and

---

<sup>154</sup> As part of this project a number of sub-grants were awarded in amounts not to exceed US\$ 275,000, to cover advisory support, training, and coaching, including related organizational expenditures. The sub-grant agreements for those training centers that had a non-public status included an obligation for these centers to also provide technical assistance to public centers in their respective area, in a later phase of the project. The Project also funded a quality control and accreditation mechanism, which was used to assess the capabilities of training centers throughout the country to deliver training modules in results-based management, PFM, procurement and human resources management in accordance with a defined set of minimum standards. (World Bank, 2011).

<sup>155</sup> The six institutions were: the National School of Administration (ENA); the National School of Finance in Kinshasa; the UNILU-Society of the University of Lubumbashi; the Salama Center in in Lubumbashi; the Catholic University of Bukavu and the University of Kisangani.



strategies. In addition, at the end of year, annual reviews of projects and programs' results were organized by UNFPA local staff in collaboration with grantees. In addition to short-term workshops and joint consultations held at the end of the year, UNFPA provided coaching to its decentralized staff towards the development of annual action plans and equipped local staff with technical monitoring forms (*fiches techniques*) that would help them in assessing the progress of the field activities against the envisaged objectives. However, such efforts were often not sufficient. As reported by a UNFPA officer:

“A monitoring culture is not there: local organizations do not monitor their activities regularly and wait until you “hit” (emphasis added by the author) them.”

In addition, as evaluation was still regarded as an external activity carried out by a third party (in most cases, a foreign firm or university), UNFPA did not invest much in enhancing CSO evaluation capacity in 2012.

### *SENAREC*

Another important national ECD stakeholder was the National Secretariat for Capacity Strengthening (*S cretariat national pour le renforcement des Capacit s* or *SENAREC* in French), established by Ministerial Order 98 of 21 February 1998, and placed under the supervision of the Ministry of Planning. As the leading governmental agency mandated to coordinate and ensure the quality of all the capacity building activities implemented by both the national government and development partners in the country, SENAREC had as a primary objective to become a one-stop window for all capacity strengthening providers operating nationwide.

Far from becoming an entity with a credentialing or licensing role, SENAREC aimed to map capacity building actors within the country by specifically defining their geographic scope and overall value of the intervention. In doing so, the intention of SENAREC was to

make sure that all the capacity building activities reviewed by them would be aligned with the national 5-year Capacity Strengthening Program (2011-2015), supported by the World Bank, that they managed. In particular, this entity managed the sub-grantee agreement with the six Centers of Excellence identified by the World Bank project.

*SERACOB*<sup>156</sup>

An NGO operating in the whole Central African region and supporting the strengthening of civil society on a variety of governance and advocacy issues, SERACOB seemed to be one of the most well placed entities within the Congolese context to provide ACoSE with technical capacity support. The organization's track record of achievements including the facilitation of participatory poverty diagnostic and analysis in communities of several DRC provinces since 2004.

*Multi-donor Pooled Fund for Capacity Building*

PRONAREC, the National Capacity Building Agency was funded by a variety of donors in the form of stand-alone earmarked support within a pooled fund scheme. The AfDB supported the development of PRONAREC's Human Resource Management capacity. AfDB also funded the development of a strategic framework for the engagement of the Congolese Diaspora (US\$ 30 million). UNDP, in contrast, allocated resources (from its larger governance budget of \$179 million) to sector-related technical assistance and sector studies. USAID instead supported civil society organizations, with a special focus on development of provincial monitoring capacity. The French Cooperation provided

---

<sup>156</sup> The acronym stands for "Service de Renforcement des Appuis aux Communautés de Base en Afrique Centrale" (Support Strengthening Service for CSO in Central Africa).

additional support (US\$ 3.4 million) to ENA as well as to the AfDB-financed human resource management capacity component (World Bank, 2011, p. 16)

*Factors Enabling the Development of An Evaluative Culture In DRC*

Based on the interviews with the current and past coordinator of the Association as well as a handful of current members conducted as part of this case study, the following list of strengths and weaknesses were identified.

*Wealth of Technical Expertise Shared by Members with the Rest of the Association*

Thanks to their professional positions in a variety of organizations (both public and private), the Association's members were constantly informed of new technical resources and national M&E related events which they could take part in;

*The Association's Ability to Provide Evaluation Consulting Services*

Through the formalization of the association and the opening of a bank account, ACoSE was able not only to receive payments for services rendered to anyone who might hire them for either conducting evaluations or reviews of the quality of the collected data (that was the case of the evaluation conducted for UNCIEF in 2011). As the ACoSE President put it in course of the interview with him:

“You can't manage the membership fees at home. Without formalization, certain activities we carried out in the name of the association (such as the evaluating of UNICEF project funded by USAID) would have not been possible.”

Similarly, formalizing the association allowed ACoSE to receive grants and others types of financial support.

### *VOPE Social and Political Legitimacy*

Besides being a driver of financial sustainability, the formalization of ACoSE status was instrumental in enhancing both the association's programmatic and organizational sustainability. ACoSE could benefit from the provisions of the National Law on Associations granting it the freedom of speech. Put simply, unlike other information associations in the DRC (e.g., the Alliance for Transparent Elections shut down in early 2012), ACoSE activities did not risk being suspended at any time based on their content;

### *Reform of the Public Service*

Despite having been put on hold since 2006, the DRC finally decided to tackle a very severe issue affecting the efficiency of its public sector, that is, the lack of a pension system for civil servants past the retirement age and the subsequent continuation of service in strategically relevant posts regardless of their absenteeism and lack of capacity in such new relevant areas as RBM and the use of modern technologies. The public service rejuvenation project, under preparation in 2011-2012, was expected to facilitate the injection of young talented staff into the Congolese public service, thus paving the way for the uptake of better M&E practices.

### *The Availability of a Central Coordinating Agency with Fiduciary Competencies*

The coordination of many activities funded under the national Congolese ECD program by a specialized capacity building agency placed under government's supervision (SENAREC) was particularly effective. In particular, such institutional arrangement contributed to the fulfillment of two primary objectives. First, the development of national ownership of programming and implementation strategies aimed at developing the national

capacity in M&E. Second, the strengthening of national training institutions (centers of excellence) through the use of well-structured sub-grant agreements.

#### Factors Discouraging the Development of an Evaluative Culture/ECD in the DRC

##### Limited Time Availability of the Most Qualified Members

Despite being very knowledgeable and resourceful, some of the ACoSE members with the most experience in evaluation had a very busy schedule and did not seem to be in a position to attend meetings or contribute to the life of the association as much as one would have hoped for. That included the limited sharing and dissemination of knowledge from the most experienced members to the less seasoned ones (that is, those very same members that decided to join ACoSE to increase their knowledge of evaluation issues and enhance their practice). As a result, members over time lost interest in participating in the association's activities due to the limited utility of the interactions.

##### Limited Availability of Both ACoSE Board of Directors and President

Despite the general expectation that the ACoSE Board members would get in touch with the membership on a regular basis by organizing a certain number of learning events on evaluation, the Board members' travel schedule prevented them from fulfilling their coordination function effectively. Due to his busy work schedule (that included being the president of a different association focused on the promotion and use of statistics for development purposes), the Association's President could not engage with both the Board and the General Assembly as he would have liked to. As the President often had to delegate some of his functions to the members within the association, a few members believed that a full-time executive secretary should be recognized in the Statute.

### Irregular Payment of Quarterly Membership Fees

Unlike some other national associations that had established yearly membership fees, ACoSE established quarterly ones. According to several ACoSE members, the frequency of the payment and the amount (over a total of US\$40 dollars per year in addition to the one-time initiation fee of US\$50) represented a hurdle for many and certainly discouraged many interested practitioners from joining the Association.

### Low Use of Prior Learning for Improvement Purposes

Some of the ACoSE members found it quite frustrating that, upon completion of an M&E training, they would get back to their office and encounter a good deal of resistance to the idea of public accountability and performance assessment. Therefore, ACoSE could play a significant role in liaising with CSOs and strengthening their understanding of, and involvement in, M&E at the national level (e.g., through their participation in Citizens' Report cards). However, due to the high implementation costs associated with such endeavors, ACoSE became rapidly aware that they could not play such mobilizing role vis-à-vis both CSOs and decision-makers without further financial support from development partners, such as UNFPA. Likewise, several ACoSE members believed that the availability of opportunities to participate in real evaluations would be a critical factor towards the enhancement of their practices and the dissemination of their lessons learned with the rest of the membership;

### Low Motivation to Participate in Association's Activities

Some of the respondents mentioned that several government officials did join the association with the belief that this would enhance their ability to participate in scoping and evaluation missions funded by their departments. Similarly, several members decided to join

as they thought that they would be receiving per diem or other monetary incentives in exchange for their participation in training programs.

The Evaluation Capacity Development Sphere-Function-Actors-Role Framework (ECD-SFAR Framework).

Based on the preliminary findings of the literature review and in line with the hypothesis the VOPEs involvement in ECD programming represented a critical contribution to the promotion of a national evaluative culture, a framework describing the main features of the ECD national ecologies was developed: the ECD Sphere-Function-Actors-Role (ECD-SFAR) Framework (Figure 7). Aimed at facilitating the targets of ECD programming and the corresponding modalities of implementation, the ECD-SFAR Framework sought to fill the gaps associated with two of the “classic” paradigms used to interpret Evaluation Capacity Development (ECD) programming in international development and discussed in Chapter II:

1. The traditional Evaluation Capacity Development “supply-demand equation” paradigm which associated the supply side with donors and the demand side with national governments; and
2. The three-level framework recognizing the relevance (but not the excessive generalization) of individuals, organizations and the enabling environment in the design and evaluation of ECD programs.

In an attempt to turn the classic ECD discourse into a more inclusive and democratic exchange on the topic of country systems’ accountability and organizational learning, this framework acknowledged the role of national and regional evaluation associations (as well as that of the academia and the private sector) in the promotion of a

national evaluative culture as well as their interactions with the other ECD stakeholders in the “within governments” “outside governments” spheres). This was aligned with some of the principles in institutional economics (e.g., social interactions influence macro-level economic activity) or modern management theory (businesses are dynamic entities and the successful sale of a product depends on a series of value-chain relations).

In addition, instead of referring to the two distinct terms of the classic ECD equation (the thinking has been for years that you are *either* on the supply side or the demand side), this framework recognizes that within each sphere (e.g., within and outside governments), each ECD actor could be *both* an evaluation consumer and an evaluation producer/initiator, based on his/her function within the system (doer or policy-maker/decision-maker). Unlike the past when actors operating “outside governments” did not receive much support from development partners, thanks to this Framework, the two spheres (within and outside governments) appeared to be finally connected with each other. Furthermore, citizens’ control was finally taken into account: the case study on South Africa was particularly enlightening in this sense as, based on it, that the more national mechanisms are in place for the population to express their voice, the stronger the influence of the public at large on the commission, use and disseminations of evaluations<sup>157</sup>.

The Framework was applied to the three case studies developed as part of this study (DRC, Niger and South Africa). As a result, it was possible to conduct a thorough assessment of VOPEs’ contribution to their respective national ECD programming discussed in Chapter 5 as well as to situate VOPE more accurately in each of the three countries’ ECD ecologies (Appendices M, N, and O).

---

<sup>157</sup> That being said, the ECD-SFAR framework was realistic enough to acknowledge the fact that citizens’ control would not be there in a number of more or less democratic contexts).



The core of the Framework was the pie (in the center). The pie included all the major ECD stakeholders (here referred to as ECD Actors) operating within a country. ECD Actors were qualified according to the three key criteria described below: Sphere, Function, and Role.

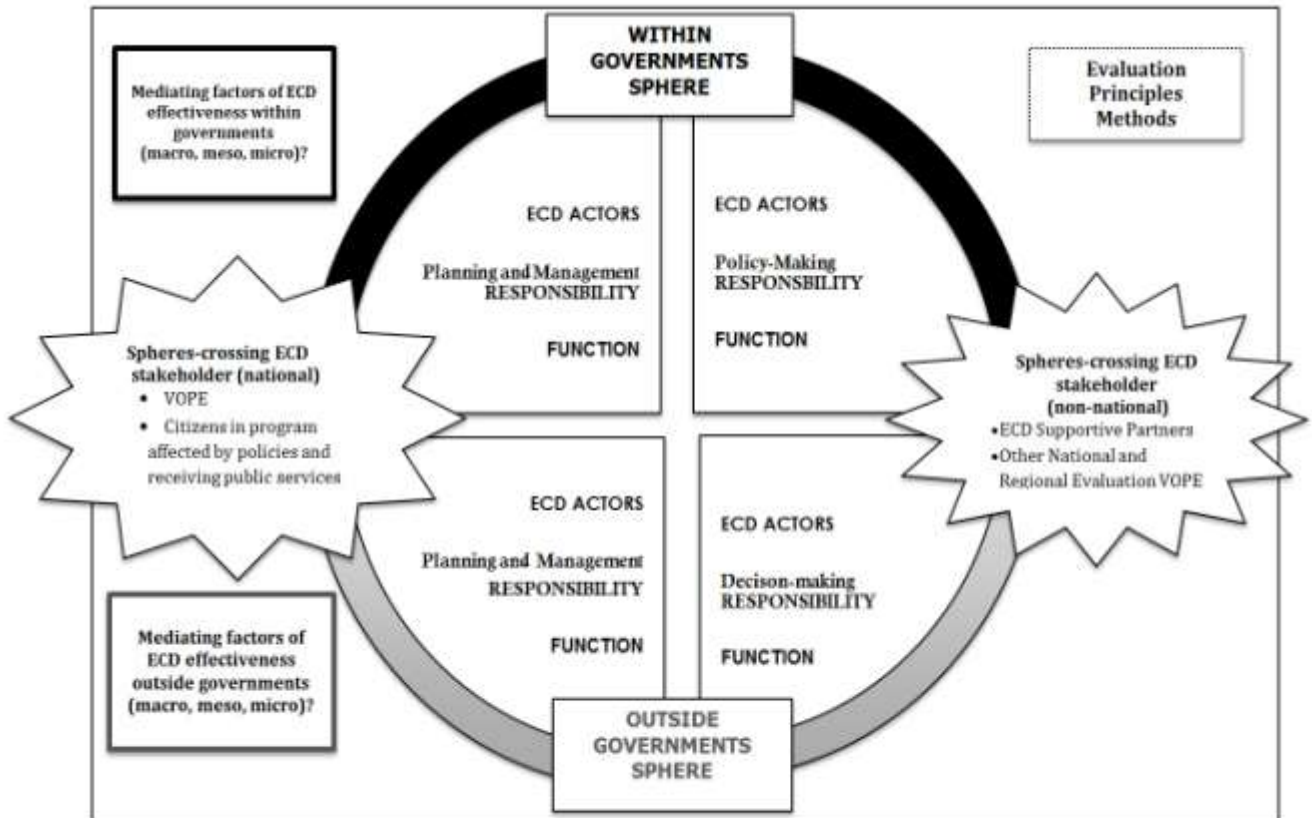


Figure 7. The SFAR Framework

## Sphere

The term Sphere<sup>158</sup> indicates the sector in which the ECD Actors carry out their evaluation functions and attend to their professional responsibilities. Since the a large part of ECB support in the past was provided to actors in the public sector, the framework includes a total of three spheres:

- a. The first sphere, including the two top slices delineated by the black colored semicircle, includes all the national ECD actors operating within the Government (Within Government Sphere);
- b. The second sphere, including the two bottom slices delineated by the grey colored semicircle, includes all the national ECD actors operating outside the Government (Outside of Government Sphere). These includes the following: Private Sector (corporate, foundations, research centers and center of excellence); Local NGO; and Institutions (Universities, Churches, sectorial, Topical and Financial)<sup>159</sup>
- c. The third sphere delineated by the two dotted stars (middle left and middle right of the pie) includes those national ECD actors (VOPE and citizens) and multi-national or non-national ECD actors (Regional Evaluation Associations and development partners, such international NGOs) that interact with both spheres (Within and Outside governments) and have the potential to facilitate the dialogue further between them.. The VOPE then become sphere-cutting ECD Actors.

---

<sup>158</sup> The circular shape was aimed at providing a more holistic and less fragmented representation of ECD, as a “human or social systems that evolve organically in unpredictable way in response to a wide range of stimuli and through multiple interactions” (Land, Hauck, and Baser 2009).

<sup>159</sup> The entities listed under private sector and institutions were drawn by Jean-Serge Quesnel’s chapter on evaluation associations and networks published in a relevant IOCE publication (Segone & Ada Campo, Eds., 2006).

## Function

The term function indicates the specific level at which ECD actors carry out their professional tasks (the level is usually suggested by each ECD actor's specific job title). The framework includes three main levels: operational level, strategic level and policy level. Each level will include a number of corresponding professional responsibilities.

-The Operational Level Actors include, among others, Planning Specialists, M&E Specialists, NGO M&E Officers, and Budget Developers (see both the top and bottom slices on the left side of the circle);

-The Decision-making level Actors include, among others, NGOs directors, NGO Board of Directors, University Provosts, heads of University Departments, CEO of private enterprises (see bottom slice on the right);

-The policy level policy-making Actors include, among others, elected or politically Appointed Government Officials (executive, legislative and Judicial branch + Public Service Commissions and Auditor General's office staff) with policy-making functions at the national and/or sub-national levels

## Role

The term role indicates the specific behavior and attitude adopted by each ED actor in relation to the evaluation activities that they pursue. This could include the following: Consumer, Provider, Initiator/Commissioner, Indirect supporter (and possibly others).

## Mediating Factors

For each of the three spheres three different levels of mediating factors that will affect ECD effectiveness:

- a) Macro-level Factors: type of political regime, type and quality of the relationship between the political majority and the opposition, history, economic environment, availability of a National Evaluation Policy, institutional locus of the evaluation function within the national administration,
- b) Meso-Level: staff Turnover, availability and accessibility of evaluation data and/or technology;
- c) Micro-Level: personal motivation, type of process adopted to select participants, years of Experience.

#### Values

For each of the four slices, the priority of the following three values was identified:

Credibility, Independence, and Quality

## CHAPTER 8

### CURRENT PRACTICES TO MEASURE ECD

#### Introduction

Based on both a systematic literature review and a series of semi-structured interviews conducted in the field with a variety of evaluation practitioners and commissioners, four sets of variables commonly used to measure ECD effectiveness in the past were identified.

The first set of variables, the most commonly used of the four presented in this section, was aimed at assessing the short-term results (such as, increase in knowledge and development of technical skills) produced by ECD activities (e.g., training, coaching, mentoring) implemented at the individual level.

The second set of variables was aimed at assessing the medium- and longer-term results produced by ECD activities implemented at the organizational and institutional levels (e.g., the adoption of mid-course corrections based on evaluation findings<sup>160</sup> or the Parliament's approval of a national evaluation policy).

---

<sup>160</sup> Henry and Mark speak of three different type of utilization: a) Individual (attitudinal change, skill acquisition, and behavioral change; b) Interpersonal (changes in justification, persuasion and social norms), c) Collective (changes in setting the agenda and modifying policy) (Henry & Mark, 2003).

The third set of variables was aimed at assessing the internal processes inherent to ECD providers<sup>161</sup> and ECD consumers<sup>162</sup>.

The fourth set of variables was aimed at assessing the relevance of contextual factors (e.g., the institutional arrangements in place within organizations being capacitated or the special political system where the ECD program was taking place) and their mediating effects on ECD outcomes.

### Variables and Levels

Based on the field interviews with a variety of ECD stakeholders as well as content analysis of the ECD specialized articles included in the systematic literature review, it became apparent that, for evaluating ECD effectiveness adequately, it was not sufficient to measure a list of pre-identified individual-level evaluative variables (e.g., change in degree of knowledge or adoption of good evaluation practices among individual evaluation officers a result of their participation in one or more ECD activities). It was also important to measure a series of organizational-level variables affecting those very same individual-level outcomes (e.g., change in evaluation practice). Otherwise said, when evaluating the effects of an ECD program on evaluation officers' knowledge and behavior, it was relevant to measure the availability and degree of effective implementation of evaluation policies (including sanctions and incentives) within those ministries or agencies where those officers worked. Therefore, for each of the four sets of evaluative variables presented in this section, the corresponding

---

<sup>161</sup> Evaluation of processes inherent to ECD providers focused, among others, on the degree of ECD targeting's inclusiveness, the level of budget resources available for implementation of activities in the field or the timeframe within which ECB activities are expected to yield their envisaged results.

<sup>162</sup> Evaluation of processes inherent to ECD consumers focused, among others, on special traits of the population whose capacity was being developed, including their needs, interests and motivation level.

level at which ECD outcome of interest needed to be measured was provided. Overall, four main levels were identified (Figure 8).

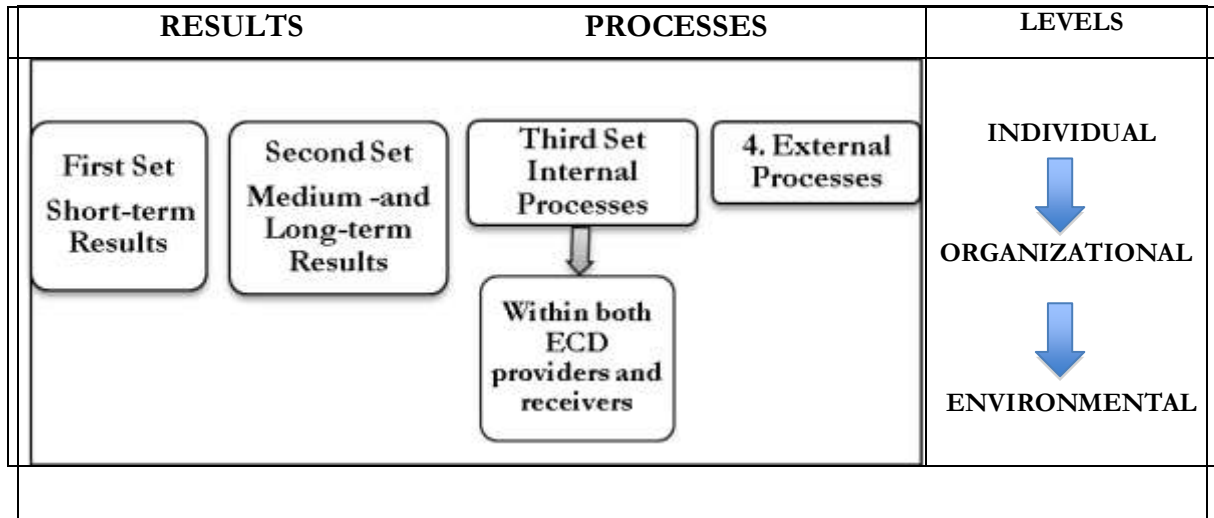


Figure 8. The Four Sets of ECB/ECD Evaluative Variables

First, individual-level variables. According to seventy-one percent (71%) (64/90) of the in-country practitioners interviewed in the field, variables at this level were focused on measuring the effects of those specific ECD activities aimed at improving the attitudes, knowledge and skills of those who were capacitated)<sup>163</sup>;

Second, organizational level variables. According to fifty-four percent (54%) (49/90) of the in-country practitioners interviewed in the course of data collection, the variables measured at this level focused on the effects of ECD activities targeting entire organizations, such as the establishment of a incentive mechanism for the use of evaluation within a ministry or the systematic use of evaluation findings for decision-making within the President’s Office.

<sup>163</sup> Institutional capacity is “a means to achieve high level programme results, rather than as an end in itself” (Lessik and Michener, 2000, p. 3).

Interestingly, according to some respondents, ECD at this level needed to be assessed through other variables that could be grouped in two different categories: institutional capital variables and technical capital variables. Institutional variables included: a) the access to incentives according to fifty-two percent (52%) (47/90) of respondents; b) the availability of a M&E Policy, according to forty-seven percent (47%) (42/90) of respondents; and c) the quality of the M&E legal framework in place including the division of evaluation tasks and responsibilities at three different levels -strategic, program management and operational according to twenty-three percent (23%) (21/90) of respondents. Technical capital variables included the availability of M&E tools and guidelines sixty-three percent (63%) (56/90), the quality of the Management Information System (38%) (34/90), and the data collection process according to twenty-seven percent (27%) (24/90)<sup>164</sup> of respondents.

Third, environmental level interventions: variables at this level were focused on measuring the changes produced by ECD on general attitudes and behaviors within society that might enhance both the national ownership and sustainability of evaluations.

#### The First Set of ECD Evaluative Variables

The first set of ECD evaluative variables was aimed at assessing the extent to which ECD activities (e.g., trainings, technical assistance and dissemination of written materials) had achieved their envisaged objectives in the short-term. According to a comparative review (Nielsen & Attström, 2011) conducted across articles written by some of the world's ECD leading scholars and published in a recent WB publication on ECD (Rist et al., 2012),

---

<sup>164</sup> In their analysis of the factors that influence evaluation activities in a organization, Gibbs et al. (2002) put forward: funding agencies' expectations, financial resources, leadership, staff (and staff stability), technologies, and the tools available for evaluation.



such variables were the most commonly measured in ECD evaluations, mostly due to the fact that trainings represented the most frequent modality of ECD programs being evaluated. Such variables commonly associated with the first two levels of the Kirkpatrick 4-Level Model (Kirkpatrick, 1998) (Table 13) were also identified as the mostly commonly used ones in ECD evaluations, as also attested by the majority of studies and evaluation reports included in the systematic review<sup>165</sup>.

Table 13

*Kirkpatrick 4-level Model: Level 1 and Level 2*

<b>Level</b>	<b>Definition</b>	<b>Measuring tools</b>
<b>Level 1</b>	Participant satisfaction	End-of-course participant questionnaires.
<b>Level 2</b>	Learning outputs	Posttests, sometimes as compared with pretests.

Source: Kirkpatrick (1998)

### The Second Set of ECD Evaluative Variables

This second set of variables was aimed at assessing the extent to which ECD activities had achieved their envisaged objectives in the medium- and long-term. Usually associated with the top two levels included in the Kirkpatrick 4-Level Model (Kirkpatrick, 1998) (Table 14), these variables appeared to be used much less frequently in ECD evaluations than those included in the first set. This finding was confirmed by a lower

---

<sup>165</sup> Similar findings were reported in a recent systematic review of ECB interventions published in the American Journal of Evaluation (Labin et al, 2011). According to the article, the most frequent individual-level outcomes reported in 93% of the studies were changes in either knowledge (51%) or practice, such as evaluation planning or implementation (80%) with a primary focus on data collection 43% (n=26), rather than in attitudes towards in evaluation. According to the same article, the strategy that had proved the most effective in leading to increased knowledge was training (n=8) whereas the one that had affected the behavior in 96% of cases (n=23) was a combination of the three different strategies (training, experiential learning and technical assistance).

percentage of studies included in the systematic review that had reported the use of such variables.

Table 14

*Kirkpatrick 4-level Model: Level 3 and 4*

Level	Definition	Measuring tools
Level 3	Performance change outcomes	Multiple, including observation, interviews and surveys of participants, colleagues, and supervisors.
Level 4	Organizational impact/results	Multiple, including comparisons with baseline organizational performance measures, surveys, and interviews with key informants

Source: Kirkpatrick (1998)

That such variables were harder to measure did not come as a surprise as medium- and long-term ECD objectives were, first, more difficult to be attained<sup>166</sup> and, second, more challenging to measure within the often-limited timeframe of ECD interventions.

With respect to Level-4 outcomes, the challenge of measuring organizational level outcomes attributable to ECD interventions appeared to be even greater due to the variety of “structural and often donor-based limitations” to the attainment of “large-scale system-level improvements and impact,” as stated by the two authors of a recent UNESCO report (Ortiz & Taylor, 2008, p. 40). That notwithstanding, a few illustrative cases of studies reporting ECD organizational-level outcomes were identified.

---

<sup>166</sup> According to a World Bank report assessing the effectiveness of trainings in strengthening national capacity, the development of organizations’ or trainees’ capacity to achieve development objectives (Kirkpatrick Levels 3 and 4) occurred only half of the time. Interestingly, based on the analysis of the findings included in the report, the under-optimal results associated with such trainings were not always due to the training providers (e.g., the lack of training modules customization to trainee’s needs and context) but were often relate to organizational factors that did not fall within the sphere of control of training providers (e.g., the lack of institutional incentives, the inadequate participants’ selection, the trainees’ poor understanding of how to use the knowledge (World Bank, 2008).

First, in Labin's systematic review (Labin et al., 2012), leadership support as well as the allocation of organizational resources for evaluation were mentioned as critical variables in 32% (20/62) and 36% (22/62) of cases respectively<sup>167</sup>. In addition, changes in culture and leadership were identified as the least common ECD outcomes.

Second, measures of evaluation culture development within organizations receiving capacity development support were included in a study focused on ECD of community-based HIV prevention programs (Gibbs et al., 2002). Aligned with the later observation (Hoole & Patterson, 2008) that many funding agencies emphasized the evaluation accountability purpose and did not focus as much they should on allowing the CSO or the ministries whom they support to develop a real learning and evaluative culture, three different ECD organizational stages of CD were identified by Gibbs and his co-authors (Gibbs et al., 2002): (i) *compliance*, (ii) *investment* and (iii) *advancement*. The *compliance* phase was typical among those organizations whose evaluation activities followed the funder's requests so closely that no adaptation or customization of the intervention to the local needs (and therefore of the corresponding M&E tools) was needed. The furthest one organization could go during this phase was to review the evaluation report internally or to conduct a data quality review (Gibbs et al., 2002). The *investment* stage was common among those organizations that had been gradually institutionalizing evaluation (e.g., by including a specific budget times line for evaluation) by also measuring medium and long-term outcomes. This was also the stage where evaluation was being increasingly used for programmatic improvement (formative evaluation). The *advancement* stage was the one characterizing those organizations which had institutionalized support for evaluation at all

---

<sup>167</sup> Interestingly, the provision of incentives for the further development of leadership support was attested by only 9.8% of the reports ( $n=6$ ). Similarly, the lack of organizational budget resources for evaluation was reported in 49% ( $n=30$ ) as the most common organizational-level barrier to evaluation.

levels among their staff and that had adopted advanced evaluation methods and design (also in response to complex evaluation)(Gibbs et al., 2002).

Third, three additional variables were identified to measure enhanced capacity: a) resilience, b) readiness and c) ability in responding to development challenges over time (Ortiz & Taylor, 2008). Fourth, a few other variables (Guijt, 2008) were identified that be used as proxies to measure changes in organizational capacity against a larger variety of learning purposes (not only evaluation related) such as: a) Financial Accountability, b) Operational Improvement, c) Strategic Readjustment, d) Contextual Understanding, e) Deepening Understanding, and f) Self-Auditing.

#### The Third Set of ECD Evaluative Variables

In this third set, a variety of evaluative variables (mostly process-related), both associated with ECD providers (e.g., the degree of targeting's inclusiveness) and ECD consumers (e.g., the size of the target population or the availability of incentives, the type and quality of interactions within the local community of practice) were included. This set of variables was quite important for several reasons.

First, assessing not only the number of people being capacitated within an organization or institution but also the ratio between those who had been capacitated and the total number of people working in those organizations and institutions, appeared to be critical. According to forty-three percent (43%) (39/90) of in-country practitioners interviewed during the fieldwork, it was not sufficient to increase the expertise of a few local evaluators to bring about changes in national evaluation capacity. It was of utmost importance that a critical mass of individuals be reached by ECD programs and that they all: a) possess an intimate knowledge of the context where the evaluation would take place

according to sixty-three percent (63%) (57/90) and b) would be given the opportunity to exchange with other more experienced evaluators or groups of evaluators, according to forty-one percent (41%) (37/90) of in-country practitioners.

Second, measuring the degree of ECD targeting's inclusiveness was quite relevant as attested by a project management professional based in Canada and highly involved in both ECD scholarly research and practice:

Selection of training participants was very unsystematic and, rather than being based on competencies, is often dictated by personal relationships between the training organizers and the individuals enrolled in the program. This entails the risk of supporting people who are more interested in institutional tourism than training contents (Clotteau et al., 2011, p. 175).

Third, assessing the availability of incentives was key to measuring ECD effectiveness. During the interview with Fred Carden's<sup>168</sup> in his IDRC Ottawa office, the importance of assessing the mediating effects of incentives as well as leadership/championship within any system where ECD is being implemented was highlighted. However, the difficulties associated with the evaluation of such ECD processes were identified, as well as, in the course of the interview. As stated by Carden:

[...] As development professionals, we need to measure well and better. In order to that, we need to open debate on how and what we measure. We also need to support and foster informed demand from multiple groups within society and break open space for citizen engagement in evaluation. We need to find mechanisms to bridge the gap between supply and demand and infuse our work with understanding of how evidence can bring change. Indigenous field building needs to be at the center of this work. [...] Of course, there is room and need for multiple voices, perspectives, debate, critique, and support from the north and south.

Interestingly, when asked about the availability of incentives made available at their office/institution to put into practice what was they had learned in the course of trainings,

---

<sup>168</sup> Fred Carden is a especially prolific author on evaluation capacity building and he is currently directing the Evaluation Unit at IDRC Canada.

sixty-seven percent (67%) (68/90) of respondents reported the lack of any incentives.

Fourth, other variables included in this set were those measuring the modalities and results of learning by doing, technical assistance, sharing practice guidelines, developing organizational policies and infrastructure (Stockdill et al., 2002). Fifth, a variety of other variables included in this set were those measuring the quality and timeliness of organizational arrangements either allowing the delivery of ECD activities or facilitating the adoption of new evaluation by the actors (both within and outside of governments) targeted by ECD programming. Indicators used to measure such variables included the following:

- Percentage of targeted ministries/provincial authorities issuing executive orders aiming to facilitate the application of the evaluation function (percentage)
- Number of certified coaches available to work with local providers of evaluation services (number)
- Percentage of local evaluation providers satisfied with the coaching support provided to them (percentage)
- Percentage of targeted units with required equipment and resources to perform key evaluation functions (percentage);
- Number of targeted units with plans to strengthen professional staff performance (1) prepared and (2) implemented;
- Type (public/private) of evaluation training providers with detailed work plans on how accredited programs will be developed (public/private);
- Number of training centers reviewed by a National Training Institution for accreditation;

- Percentage of institutional improvement plan targets achieved within the agreed timeframe (percentage);
- Number and percentage of evaluation contracts signed by targeted units through open competitive process

### The Fourth Set of ECD Evaluative Variables

The variables included in this last set were those measuring the mediating effects of contextual factors, such as the availability of evaluation-related legislation and the presence of a favorable political environment that could both promote and sustain evaluation (Durlak & duPre, 2008; Patton 2011; Wubneh, 2003)<sup>169</sup>.

In addition, variables assessing the state of the national evaluation culture were included in this set. A very difficult construct to measure, evaluation culture could be assessed in a variety of ways. However, one of the tools identified in the course of the literature review that seemed particularly suitable to gauge evaluation culture in any given context, was a list of nine questions developed by a former senior officer of the Swedish National Audit Office in the early 2000s (Furubo et al., 2002):

1. Does evaluation takes place in many policy domains?
2. Is there a supply of evaluators specializing in different disciplines who have mastered different evaluation methods, and who conduct evaluations
3. Is there a national discourse concerning evaluation in which more general

---

<sup>169</sup> Wubneh conducted an evaluation of the African Capacity Building Foundation (ACBF), a joint initiative of African governments, the World Bank, the United Nations Development Program, and the African Development Bank. In her work four major factors affecting the successful completion of ECB outcomes were highlighted: the influence of political and institutional factors on capacity-building projects; the relationship between human and material resources and capacity building; the relative unimportance of donor funding and expatriate staff; and the importance of contextual factors such as the project gestation period, project structure, and institutional setting (Wubneh, 2003).

discussions are adjusted to the specific national environment?

4. Is there a profession with its own societies or frequent attendance at meetings of international societies and at least some discussion concerning the norms and ethics of the profession?
5. Are there institutional arrangements in the government for conducting evaluations and disseminating their results to decision makers?
6. Are institutional arrangements present in Parliament [or other legislative bodies] for conducting evaluations and disseminating them to decision makers?
7. Does an element of pluralism exist: that is, within each policy domain there are different people or agencies commissioning and performing evaluations?
8. Do evaluation activities take place within the Supreme Audit Institution?
9. Is the evaluation done focused only on the relation between inputs/outputs and technical production or also on program or policy outcomes as their object (especially policy evaluations)?

#### Are All ECD Effects Measurable?

Given the vast array of activities implemented as part of ECD interventions and the impossibility of tracking all national government or CSO networks decisions affected by ECD, not all ECD effects seemed to be measurable. Such thought was echoed in the written response of developed by the World Bank Institute (WBI) Management Team to the critical IEG report on the effectiveness of WB funded-trainings over the last decade (World



Bank, 2008). According to the managerial feedback to the conclusions of the report in question, it was quite unrealistic to use tangible and monitorable training results (based on levels 1-3 of the Kirkpatrick Model) as the indication of capacity building success.

According to the report, the challenge of assessing training effectiveness was even greater when the trainings being evaluated has been aimed at providing clients with policy options that would often translate in concrete actions only in the long-term.<sup>170</sup>

The challenges of evaluating ECD effectiveness was also confirmed by the stakeholders interviewed during the field data collection. Sixty-seven percent (67%) (60/90) of in-country participants stressed the challenge of measuring the impact of ECD work. When asked how ECD get measured, seventy-two percent (72%) (64/90) of respondents mentioned that a list of questions on what participants liked the most and the least about the training was usually distributed at the end of the training, while twenty-seven percent (27%) (24/90) mentioned that a list of more technical questions were asked before and after the training as a way to gauge participants' learning. Thirty-one percent (31%) (27/90) of in-country practitioners and funders of ECD programs interviewed during data collection also mentioned the use of case study while another sixty-seven percent (67%) (61/90) of respondents mentioned self-assessments or written questionnaire.

Furthermore, twenty-four percent (24%) (22/90) of respondents reported that it was impossible to evaluate ECD effect effectively due to three main reasons:

1. The complexity of activities implemented as part of ECD programs, according to fifty-four percent (54%) (49/90) of respondents;

---

<sup>170</sup> That was especially true to non-lending services provided within the scope of the Long-Term Strategy Exercise undertaken by the Bank).

2. The gap between the long time needed for results to come about and the shorter projects life-cycles, according to sixty-two percent (62%) (56/90) of respondents; and
3. The large variety of stakeholders involved, according to forty-three percent (43%) (39/90) of respondents.

According to the in-country practitioners interviewed in the course of field data collection, such challenges would be overcome in three different ways: allocating more resources to conduct evaluations of ECD processes, according to sixty-two percent (62%) of respondents (55/90)<sup>171</sup>; harmonizing donors' implementation efforts and reporting requirements, according to fifty-seven percent (57%) (52/90)<sup>172</sup>; and implementing follow-up assessments more systematically after the end of ECD projects, according to fifty-one percent (51%) (46/90) of respondents.

When asked what some new or emerging evaluation methodologies that they thought would be particularly effective to capture the effect of ECD programs, respondents suggested several ones, including: network Analysis, according to forty percent (40%) (36/90) of respondents; outcome mapping, according to thirty-percent (30%) (29/90) of respondents; contribution analysis among ECD funders according to thirty-one percent (31%) of respondents (28/90); the Most Significant Change methodology, according to twenty percent (20%) (18/90) of respondents; Pre- and Post- Environmental Scan, according to ten percent (10%) of respondents (9/90); "Institutional analysis," according to

---

<sup>171</sup> The availability of budget resources was identified as one of the prerequisites of effective ECD programs. Aware of the financial implications of implementing ECD programs, in-country practitioners, when asked about the greatest shortcomings associated with ECD, mentioned the corresponding high implementation costs 42% (*n*=46).

<sup>172</sup> Interestingly, when asked about the feasibility of donors' harmonization on ECD programming, 34% (*n*=30) of respondents indicated that getting donors to harmonize their ECD efforts was easier said than done.

seven percent (7%) (7/90) of respondents; “Drivers of change” analysis , according to two percent (2/90) ; and Write-shops, according to one percent (1%) (1/90) of respondents..

#### Difficult Comparability across ECD Interventions

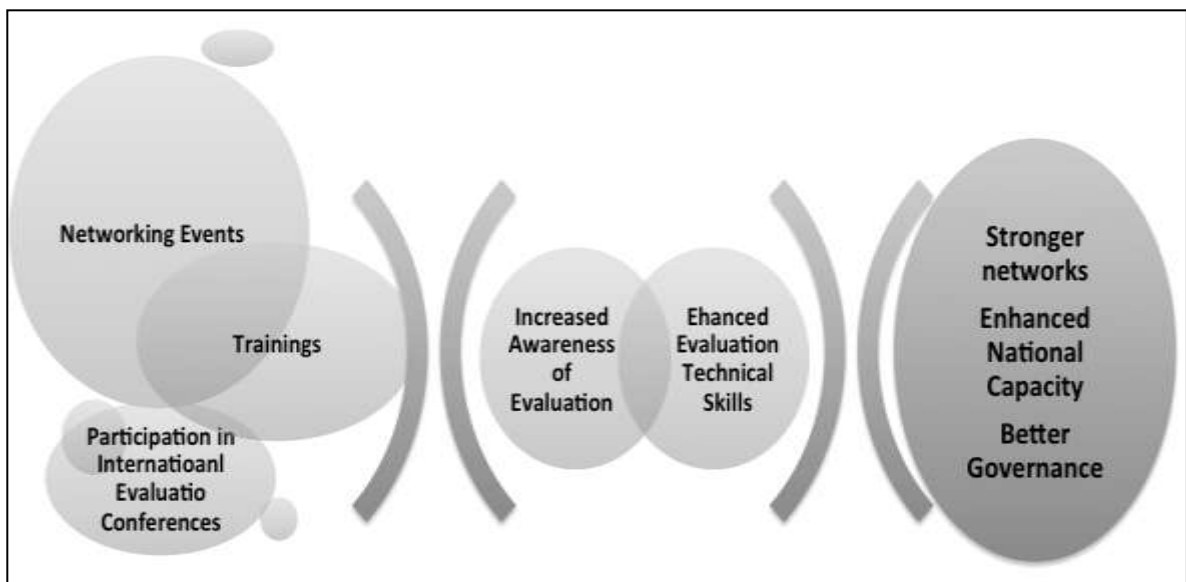
Given the plethora of initiatives funded as part of large-scale ECD programs and due to the tailoring of ECD initiatives to contextual needs, comparing the effectiveness of activities across ECD interventions was particularly challenging. That notwithstanding, there appeared to be an increasing effort to standardize some of the tools used in the course of ECD implementation. According to a representative of the South Asia Community of Evaluators, for instance, the standardization of programs was a very helpful expedient that ought to be pursued in the future to enhance a more credible comparison across similar ECD programs in the region.

#### VOPEs Evaluation

As the opportunities for further involvement of VOPEs in the implementation of more effective ECD programming in the future was one of the key areas of focus of this dissertation, a review was conducted of VOPEs evaluations conducted in the past, both within and outside the three countries where the fieldwork was collected. Interestingly, while a series of very well articulated and informative case studies (Holvoet et al., 2011; (OCE/Eval Partners, 2012) were identified, no specific VOPE evaluative framework was found in any of them.

While the need to capture the uniqueness of each VOPE through a case study was justified in the past, the wealth of knowledge built around VOPEs experiences and dynamics over the last year lend itself to some further analysis and systematization. Therefore, in an effort to contribute to future VOPEs evaluative efforts and based on the data collected in

the DRC, Niger, and South Africa, a VOPEs theory of change was developed. The theories of change presented in this section were aimed at addressing the apparent lack of logical causal pathways in the design and implementation of VOPEs activities (Figure 9). What was especially observed in the course of the fieldwork was the lack of direct links between the VOPEs missions and objectives and the variety of activities often implemented without a clear strategy by the VOPEs themselves.



*Figure 9. Apparent Lack of Logical Causal Pathways*

Despite the intention to contribute to a better understanding of VOPEs dynamics so as to enhance the planning and funding of VOPEs activities in the future, the theories of change presented in this section were meant to be a purely exploratory effort in that they did not claim any generalizability until further discussions with a larger number of VOPES around the world would be held in the future.

## First VOPEs Outcome: Knowledge Building

Based on the fieldwork, it was apparent that VOPEs would be instrumental in both the sharing and dissemination of evaluation information among its members as well as in the building of local knowledge on evaluation methods and practices (Figure 10).

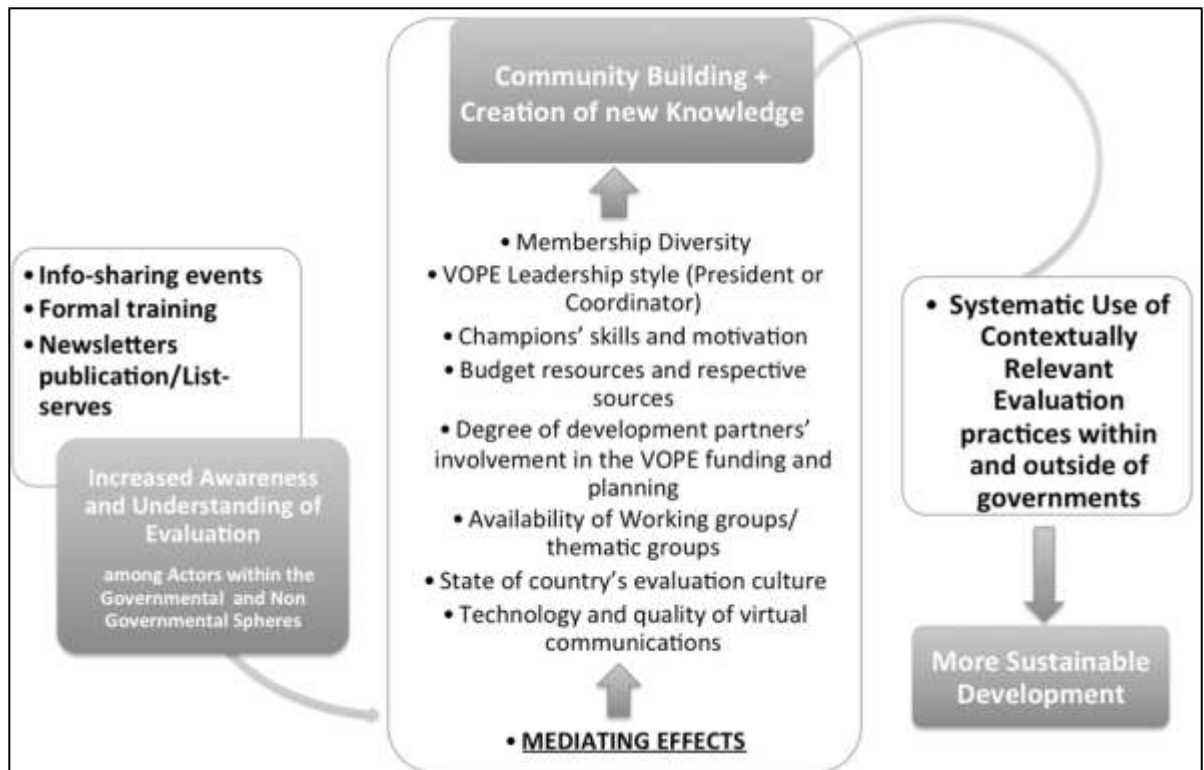


Figure 10. Theory of Change Behind First VOPEs Outcome

## Second VOPEs Objective: Professionalization of In-country Evaluation Practices

By raising the evaluation practice national standards, VOPEs appeared to be key in the professionalization of national evaluation professionals: both commissioners/managers of evaluation, and actual evaluation practitioners (Figure 11).

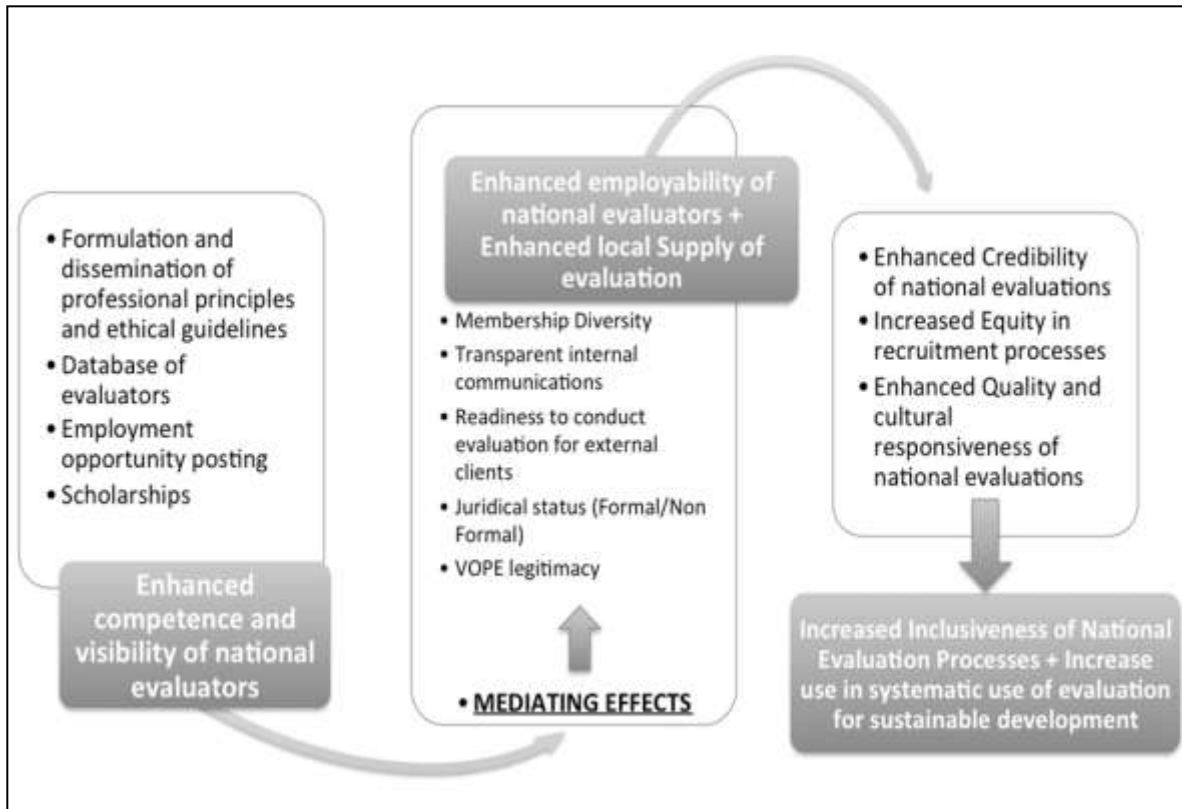


Figure 11. Theory of Change Behind Second VOPEs Outcome

### Third VOPEs Objective: Enhancing Institutional Relations

Finally, VOPEs appeared to be particularly successful in developing relationships (both formal and informal) with a variety of national and international institutions. In particular, VOPEs provided in the DRC, Niger and South Africa, a common platform for discussion on evaluation issues aimed at enhancing national practices and policies in this area. More importantly, due to the variety of their membership, VOPEs allowed actors who operated in different spheres (within and outside national governments) to interact with each other and reach a common understanding of evaluation practice and purposes (Figure 12).

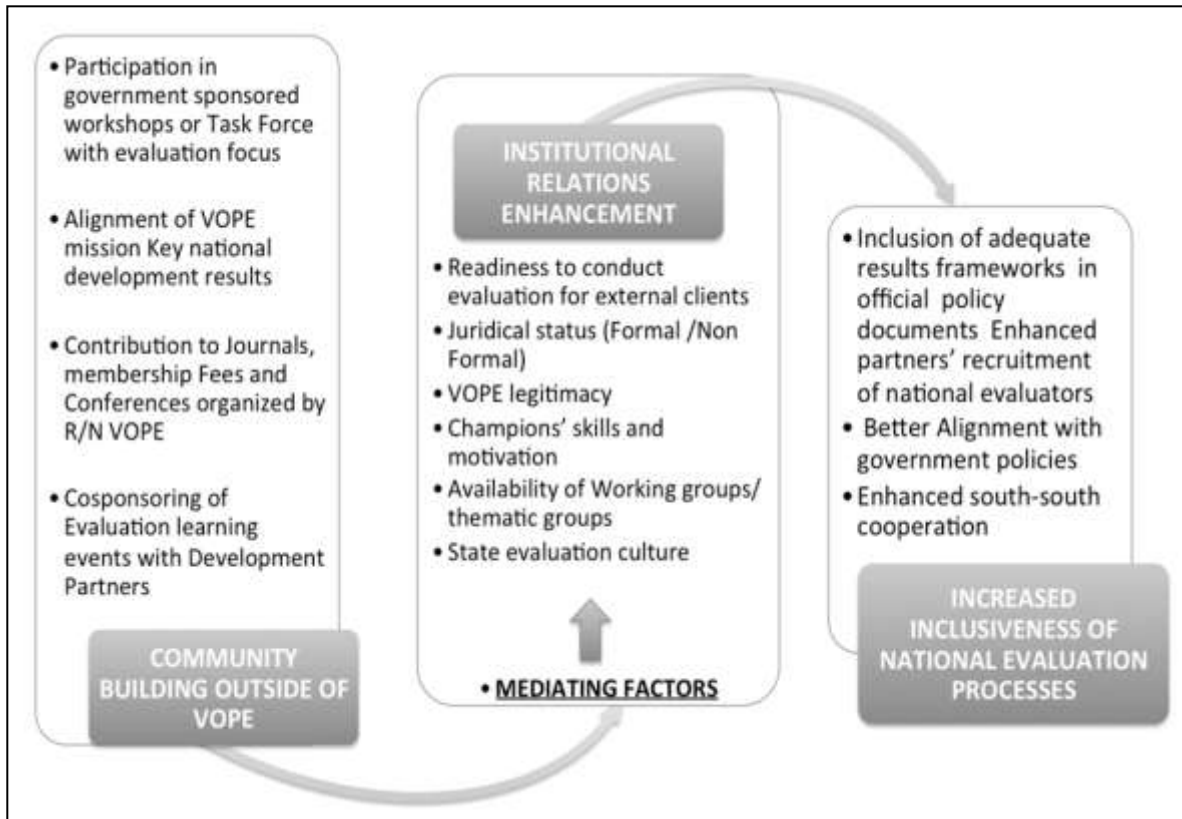


Figure 12. Theory of Change Behind Third VOPEs Outcome

### Measuring VOPEs Organizational Development

As the evaluation of organizational processes emerged as critical to the assessment of VOPEs results more effectively, in the future, the need for a tool that could measure VOPEs organizational dynamics was identified before the start of data collection. The instrument used in the course of data collection (referred to this study as the VOPEs Capability Assessment Tool), that had been adapted by a more general Capability Assessment Tool developed by ECDPM in 2008, proved to be quite effective. However, based on the learning occurred in the course of this study (e.g., new emerging dimensions of VOPEs organizational dynamics), the tool in question was slightly revised.

## CHAPTER V

### CONCLUSIONS

This last chapter consists of four sections. In the first section, the answers to the three key research questions (What are the typical attributes of ECD and what is its relationship with ECB? To what extent could ECD targeting become more inclusive in the future? How can, or should, ECD be best evaluated?) are provided. In the second section, the limitations of the study are discussed. In the third section, the study's potential implications for ECD theory, practice and programming; are presented. In the fourth and last section, prospective areas of future research are enumerated.

#### First Research Question

1. To what extent is Evaluation Capacity Development (ECD) distinct from Evaluation Capacity Building (ECB) in international development contexts?

1.1 What are the central attributes of ECD as compared to those of ECB?

Given the paucity of peer-reviewed literature on ECD, the interchangeable use of the two terms in grey literature and the ideological distinction held between ECB and ECD, the need for a better understanding of what Evaluation Capacity Development (ECD) was and how it related to Evaluation Capacity Building (ECB) was identified. As a result of the semi-structured interviews conducted with over one hundred ECB and ECD practitioners around the world, and based on the analysis of three case studies developed in the Democratic



Republic of Congo, Niger and South Africa during the first semester of 2012, the following new definition of ECB was developed:

ECB was defined as:

A necessary (but not sufficient) condition for ECD to take place. ECB mainly consists of a vast array of trainings and coaching activities (some of which are short-term in nature) aimed at building capacity, especially where capacity is either very low or thought not be in place yet, among a discrete number of individuals working either *for* or *within* organizations and/or institutions that develop, commission, manage, conduct and/or use evaluation (Figures 14, 15).

Although it is an integral component of most national and international development projects today, ECB has often been viewed (especially outside of the US) as a relatively limited accountability-driven tactic rather than a full-fledged strategy aimed at attaining organizational learning as well as other developmental objectives. As a result, ECB scope and modalities of delivery have often been considered too narrow and its goals have been described generally as being more limited in scope than those pursued by ECD.

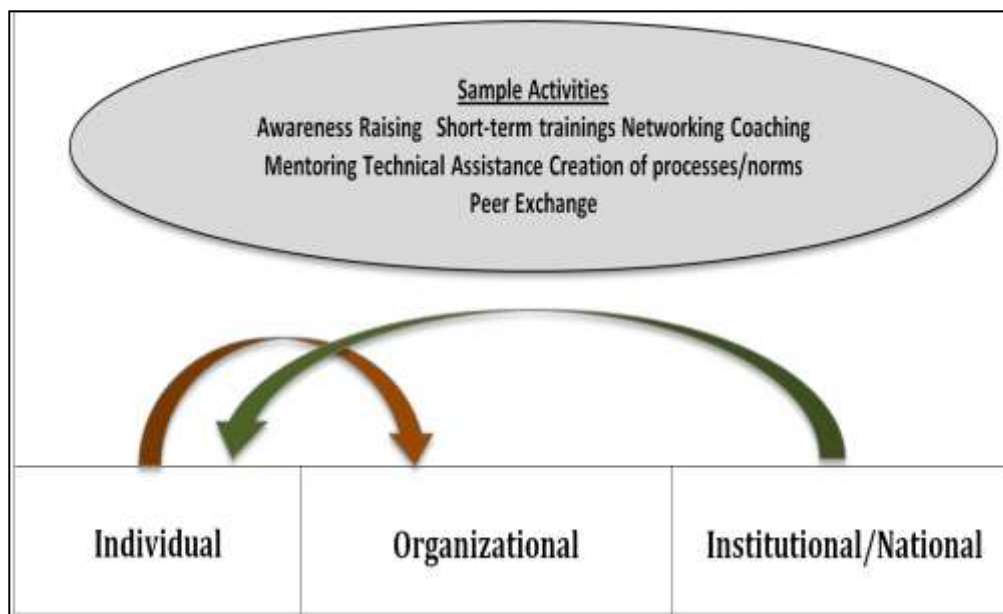


Figure 13. A Simple ECB Programming

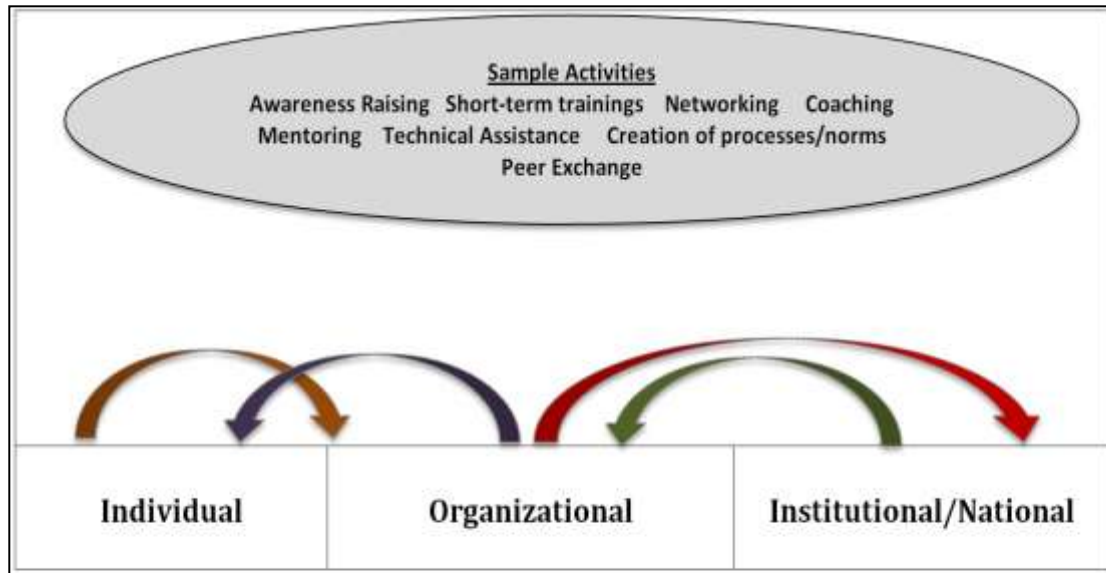


Figure 14. A More Comprehensive ECB Programming

Conversely, ECD was defined as:

A process consisting in both the integrated enhancement and maintenance over time of

- Individuals' knowledge, skills and attitudes;
- Organizations' capabilities, and
- Institutions' readiness,

Towards contextually relevant planning, management, implementation, and use of evaluation at any level-global, regional, national or sub-national.

More specifically, ECD is aimed at both individual and collective transformational learning in the pursuit of three primary goals: strengthening the technical quality and ownership of national evaluation processes; enhancing the local authenticity and cultural appropriateness of evaluation approaches, methods and tools used in-country; and increasing the use of evaluation findings as a way to improve development interventions in a variety of sectors.

In order for ECD to be successful, it is critical that ECD-savvy strategies be implemented either in a simultaneous or intentionally sequenced fashion (Figure 16). ECD-savvy strategies are specifically aimed at promoting conditions that support ECD among a variety of *actors* operating in two different *spheres* (within and outside of national government) and characterized by different *functions* (operational and policy- or decision-making) and *roles* (consumers and providers of evaluation). ECD strategies consist of a combination of short-, medium- and long-term activities (including training, mentoring, coaching, peer exchange, the creation of evaluation units or the

development of evaluation norms and policies) implemented in response to the identified actors' and institutions' needs and interests.

Moreover, in an effort to enhance ECD ownership and sustainability, it is relevant that sphere-crossing entities, either national (e.g., Voluntary Organizations Promoting Evaluation or VOPEs) or multi-national (regional evaluation associations), actively contribute to ECD-savvy strategies. Such sphere-crossing entities are especially valuable as they may serve not only as brokers of different emerging needs and interests among national ECD stakeholders but also as promoters of improved and innovative evaluation knowledge and practices among the actors animating the ECD process in the different spheres. In implementing ECD-savvy strategies, it is also critical to act according to such principles as equity, independence, quality, respect and transparency.

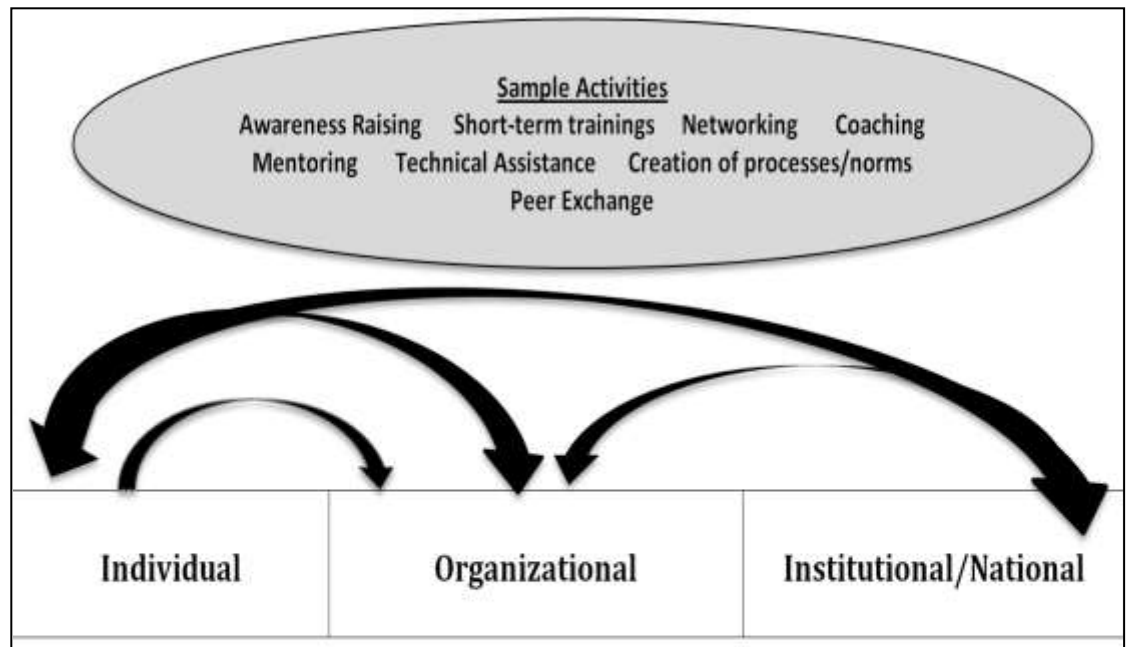


Figure 15. A True ECD Programming

Overall, as suggested by the results of the interviews held with scholars and in-country practitioners presented in this study, ECD appeared to be a systemic and adaptive process rather than the combination of stand-alone activities aimed at enhancing capacity at either the individual, organizational, or institutional level. As stressed by the definition provided above, ECD emerged as an endogenous process *building upon* (as opposed to

*building from scratch*) the existing levels of knowledge, skills and attitude (individual), capabilities (organizational) and readiness (institutional) *simultaneously or sequentially* and *in a variety of contexts* (global, regional, national, and sub-national).

In addition, the ECD primary focus identified in this study was not on either activities or products (e.g., the conduct of a training on mixed methods or the timely submission of a mid-term review to the project funder) but rather on processes, interactions, incentives, leadership, organizational learning, and organizational development.

Furthermore, ECD was characterized as a particularly inclusive process that was able to respond to continually emerging needs and interests not simply of individuals, organizations or institutions; but rather of individuals *situated within* organizations and institutions interacting with each other both in the governmental and non-governmental spheres.

Likewise, rather than resting upon a linear and mechanistic planning, ECD was understood and defined as a process grounded on both a realistic understanding of the world's complexity and the need for adopting more flexible and iterating planning processes.

To the contrary, ECB was mainly described as an activity focused at developing *separately and not simultaneously* individual technical skills (e.g., how to write Terms of Reference or develop a sampling strategy), organization-level capabilities and institutional readiness based on relatively linear and result-based planning processes. Furthermore, in contrast to the peer-reviewed literature definition of CD as an incessant endogenous process, the term ECB was described as resting on the main assumption that in-country capacity was static (you either have it or you don't) and that targeted interventions, often funded from external development partners, were the most effective in making it develop. As a result, ECB did seem to fail to recognize the institutional processes and social dynamics

in continued evolution, such as those associated with national planning processes, that had brought about change and injected innovative ideas into the system<sup>173</sup>. Similarly, a large number of respondents stated that the majority of evaluation workshops or trainings that they had taken part in lasted only for a couple of days and often consisted of a rather broad presentation of basic M&E concepts of dubious applicability to their respective work.

### Going Beyond the Existing Paradigms

Based on this new understanding of the relationship between ECB and ECD, some of the most popular related theoretical frameworks in use over the last decade were analyzed. Based on the identification and analysis of the main weaknesses associated with such frameworks, a number of recommendations were developed in order to frame the future ECB and ECD discourse in more innovative terms. With respect to the first of the three paradigms (Boyle, Lemaire, & Guerrero; 1999), two main weaknesses were identified. First, the lack of an in-depth stakeholders' analysis: ECB stakeholders were depicted generically as if they belonged to either the donors or the aid recipients group. Second, the rigidity of the role and responsibilities assigned to each side of what was referred to as the ECB equation (countries being capacitated were depicted as consumers whereas aid agencies were described as providers of evaluation funding and services). The recommendation put forward to the gaps observed in relation to this first theoretical framework was double-fold. On the one hand, it was suggested that ECB and ECD stakeholders be defined more precisely and inclusively (e.g., taking VOPEs into account). On the other hand, ECB and

---

<sup>173</sup> Of all the ideas exchanged with the large variety of respondents met during field data collection, one that seemed to have captured more effectively than others the level of appreciation for the current state of CB was the following; “CB is nothing more than parachuting evaluation trainers or organizing short-term M&E workshops in countries, to serve only one purpose despite the good intention: to make donors happy.”

ECD planners were encouraged to warrant the fluidity of evaluation roles and functions across ECD actors (government officers who developed evaluation Terms of Reference and commissioned evaluations were no longer to be seen simply as consumers but also providers of evaluation in that they could also provide evaluation trainings or coaching sessions to other colleagues within their units).

With respect to the second theoretical framework (Preskill & Boyle, 2008), the main identified weakness was the relatively limited definition of Communities of Practice (CoP): a VOPE was described merely as an “organizational-level ECB strategy” to pursue (Preskill & Boyle, 2008)—rather than a key stakeholder group to involve in the planning and implementation of an ECD program. The recommendation put forward to address the shortcomings of this second theoretical framework was to qualify CoP (especially those with a more diverse membership and a broader scope of activity) as a group of people possessing skills and capabilities that might be tapped in for the sake of planning and implementing ECD-savvy strategies.

As far as the third framework (Heider, 2011; UNDP, 2011) is concerned, its main contributions were first identified. Often summarized by a three-level diagram including enabling environment at the top, organizations in the middle and individuals at the bottom), this framework was particularly commended for having advanced the ECD discourse over the last decades in three different ways. First, by explaining the influence of current national policies, legislation, power relations and social norms (enabling environment) on ECD processes. Second, by describing the repercussions of financial and human resources as well as internal policies and procedures (organization-level issues) on organizations’ involvement in ECD. Third, by highlighting the extent to which in-country practitioners’ leadership, knowledge, skills, attitudes and practices in evaluation (individual-level issues) contributed to

the acceptance and further development of evaluation. However, two main limitations were identified in relation to this theoretical framework. First, the specific details on who the stakeholders are that ought to be targeted at each of the three levels by ECD were not specified. Second, the interactions existing across the three levels, both in terms of gaps and opportunities for collaboration, as well as the corresponding processes, were not characterized precisely enough. The recommendations put forward to address these two limitations were, first, to provide more details on the specific actors (for each of the three identified levels) whom it would be critical to get involved in ECD programming; and second, to understand the inter-level processes and to better integrate the ECD strategies not only across the three levels but also within each of the three levels.

Far from being the result of a skillful expression of purely intellectual rhetoric, the recommendations put forward in this study were intended to inform, inspire and underpin a few radical changes in the mindset and practices of ECD and ECB professionals (both among practitioners and funders)<sup>174</sup>. As a result, two specific epistemological paradigms (systems thinking and complexity) as well as their adaptation to specific ECD evaluation needs was particularly advocated for in this study. On the one hand, the use of the systems-thinking paradigm in ECD programming was recommended as it would allow identifying the individual perspectives that ECD actors have about their roles and responsibilities within the specific ECD context where they operate. On the other hand, the adoption of a complexity lens in ECD was encouraged as this would facilitate the understanding of the intimate and interdependent relationship existing among individual elements within the system and

---

<sup>174</sup> As a result, the conclusions of this dissertation would be better appreciated if understood at a meta-level.

between these and the context in which they are embedded<sup>175</sup>, especially during the ECD planning and vision-building phases<sup>176</sup>.

## 1.2. What, if any, is the relationship between the terms ECB and ECD?

The hypothesis that ECD did not completely overlap with ECB, with respect to its premises, focus and objectives, seemed to be supported by this study. Initially thought to be two opposite concepts (binary perspective), ECB and ECD ended up emerging as complementary constructs (holistic perspective). In particular, ECB was identified as the combination of interventions aimed at strengthening capacity at some but not all three levels (individual, organizations and institutional) of what was referred to in Chapter IV as the “ECD continuum” (see Appendix F).

As depicted by their visual comparison across a variety of criteria, ECB and ECD appeared to be quite contingent rather than opposite terms. Based on the analysis of the relationship between ECB and ECD, it was concluded that a two- or three-day workshop aimed at in-country evaluation practitioners and primarily focused on sampling methodology or evaluation ToR development, could be interpreted in two different ways. It could be regarded as a “business as usual” individual-level ECB activity or as the integral component of a larger ECD program provided that four critical conditions were met.

---

<sup>175</sup> At a time when impact evaluations’ aim was at identifying the causal link between a given intervention and a specific outcome of interest, under the assumed plausibility of such inferences due to the use of a counterfactual, this study was aligned with the idea that measuring attribution between an independent and a dependent variables while holding all other things constant, is not as straightforward as it might appear at first. As ECD practitioners are called upon to increasingly take feedback loops, emergent properties and interconnectedness among stakeholders operating within the system into account, Randomized Controlled Trials do not seem to be the ideal solution (Vis-à-vis the uncertainty of the broader context or enabling (or hindering) environment where ECD interventions take place, ECD practitioners will need to learn from failures and unexpected outcomes. ODI 2008, p. 25).

<sup>176</sup> Far from pursuing linear and apparently logical and infallible planning strategies, this study showed that “flexible adaptiveness” is needed and that a sound diagnostic (a sort of ECD environment scan and a continual monitoring of all the forces impacting the outcomes of ECD interventions are key to development effectiveness.



First, it was necessary that the training be intentionally and explicitly (not just nominally) linked with a variety of other activities and programs aimed at enhancing evaluation capacity among a broader number and types of stakeholders, not only at the individual but also at the organizational and institutional levels.

Second, it was of utmost importance that the training be provided by a national training institution (adequately complemented by special external support—better if from the region—in case of weak capacity locally available).

Third, it was critical that the training be developed based on the results of a participatory capacity needs assessment, as also verified by participants' high ratings of the training relevance.

Fourth, there was a need for the training to be flexible enough to respond to emerging needs (unexpressed and unexpected at the time the needs assessment was conducted) among the targets to be capacitated.

Based on the distinction between ECB and ECD and in order to strengthen the link between the two, several strategies were identified in order to make ECB more effective. First, it was concluded that the longer ECB activities were implemented at each one of the three levels (individual, organizational and institutional), the more likely they were to attain their specific objective. Second, the more levels at which capacity strengthening strategies were implemented simultaneously or as part of a well articulated strategy, the more effective they were likely to be. Third, the more integrated activities were across levels and ECD planners (both national and non-national), the more effective they were likely to be. Fourth,

the more creative<sup>177</sup> and adaptive (ECD-savvy) to emerging and unexpected scenarios ECB implementers proved to be, the more successful their efforts were likely to be.

## Second Research Question

To what extent could ECD targeting in international development contexts become more inclusive in the future?

### 2.1 What is the current capacity of VOPEs?

Based on the results of the capability assessments conducted among the three VOPEs in the DRC, Niger and South Africa, it appeared that their ability to fulfill their respective missions and objectives could not be fully understood without taking a variety of factors into account. Once identified in the course of the study, such factors were grouped in two different categories: organizational processes and organizational capabilities.

#### Organizational Processes

The variables in this category, regarded as critical determinants of VOPEs' success in fulfilling their missions, included the following:

- a) Historical Development (informal development, formalization, expansion, regional consolidation, transition to national ownership, stagnation/implosion, self-reflection, nominal/effective revitalization)
- b) Organizational Development Phase (pioneer, differentiated, integrated) (Ubels et al, 2010)
- c) Membership Diversity<sup>178</sup> (low, medium, high)

---

<sup>177</sup> That would include the application of theater principles: ECB implementers could play the role of a playwright, director, actor, or audience (Porter, 2011; Schein, 2009).

- d) Compliance with internal government rules<sup>179</sup>(low, medium, high)
- e) Degree of Internal Networking (low, medium, high)
- f) Resilience (low, medium, high)
- g) Leadership Type (concentrated, decentralized, shared) (Ubels et al., 2010)
- h) Ownership (low, medium, high)

Once identified, such variables were used to assess VOPE capabilities in all three countries (Table 15).

Table 15

*VOPEs Organizational Processes: A Comparative Checklist*

Country	DRC	Niger	South Africa
Criteria			
<b>Historical Development</b>	Stagnation	Nominal	(Expansion and
<b>Organizational Development</b>	Differentiated	Revitalization	Regional Consolidation)
<b>Phase</b>		Differentiated	Integrated
<b>Membership Diversity</b>	**	***	***
<b>Compliance with internal government rules</b>	*	**	***
<b>Degree of Internal Networking</b>	*	**	**
<b>Resilience</b>	*	**	***
<b>Leadership Type</b>	Concentrated	Concentrated	Decentralized
<b>Ownership</b>	*	**	***

<sup>178</sup> Diversity among VOPEs' regular members as well as within the Coordinating team was assessed at two different levels: a) functions (operational, decision-making, policy-making); and b) spheres of occupation (within or outside national government).

<sup>179</sup> This would include the extent of meetings regularity (both in-person and virtual) among its members.

## Organizational Capabilities

The Capabilities included in this category, considered to be particularly relevant to understand how VOPEs function and operate, included the following:

- a) Capability to Cross Spheres (sphere bounded, sphere border approaching, sphere crossing)
- b) Capability to Commit and Act
- c) Capability to Generate Development Results
- d) Capability to Relate
- e) Capability to Adapt
- f) Capability to Integrate

Once identified, such variables were used to assess VOPE capabilities in all three countries (Table 16).

Table 16

### *VOPEs' Organizational Capabilities: A Comparative Summary*

Country	DRC	Niger	South Africa
<b>Capabilities</b>			
<b>Cross Spheres</b>	*	***	**
<b>Commit and Act</b>	*	**	***
<b>Generate Development Results</b>	*	*	**
<b>Relate</b>	*	**	***
<b>Adapt</b>	*	**	***
<b>Integrate</b>	*	*	**

Overall, of the three VOPEs whose capabilities were assessed in this study, the one in South Africa (SAMEA) seemed to be the strongest one, more so in terms of capability to commit and act, relate and adapt. Besides the well-developed web of formal and

informal partnerships with governmental agencies, universities and regional VOPEs, SAMEA proved to be particularly skillful in using social media for enhancing not only info sharing but also knowledge building on evaluation among its members. As a result, despite the lack of a formal credentialing mechanism in South Africa, SAMEA seemed to have contributed to the promotion of evaluators' professionalization nationwide.

The VOPE in Niger (ReNSE), too, was able to put in place a series of activities aimed at giving evaluation more visibility nationwide and it developed a very close link with some development partners (especially UNICEF and UNDP). ReNSE members were able to provide an indirect contribution to national development policies through the participation of some of its members in both the PRSP planning processes and the drafting of a national evaluation strategy.

The VOPE in the DRC (ACoSE) emerged, among the three, as the one with the lowest capability, mostly due the paucity of activities funded by the Association as well as two other factors: the lack of compliance with internal government provisions and the relatively limited interactions with other national agencies, development partners and regional VOPEs.

2.2. To what extent are VOPEs currently involved in the conduct of evaluations and the promotion of an evaluative culture in their respective countries?

Of the three VOPE assessed in this study, only the one in the DRC (ACoSE) was able to conduct an evaluation as a collective entity. ReNSE members expressed interest in conducting an evaluation, provided that part of the revenues would be shared with the network, but they never had such opportunity. SAMEA instead expressed a strong

opposition to the idea of conducting evaluations, mostly based on its willingness to comply with its non-profit status.

That notwithstanding, all three VOPE contributed to the promotion of an evaluation culture in their respective countries by pursuing three main objectives, as highlighted in the three theories of change developed retrospectively and presented in Chapter IV:

- a) Enhancing info-sharing and knowledge building on evaluation both within and outside their membership,
- b) Raising national standards of practice among evaluation professionals, and
- c) Enabling relationships with a variety of institutions that have a vested interest in evaluation nationwide.

More precisely, in order to fulfill these three objectives, VOPEs implemented a variety of activities, the most frequent and popular of which were: to play a policy advisory support role, to conduct meetings with other institutions interested in evaluation, to organize trainings and educational opportunities for both members and non-members, and to advocate for the more frequent implementation of evaluation and use of the corresponding findings.

Overall, based on the retrospective formulation of a VOPE theory of change and the identification of four main outcomes associated with their activities, four main VOPEs contributions to the promotion of an evaluative culture were identified:

- a) Knowledge Sharing (discontinued, episodic, systematic)
- b) Knowledge Building (absent, episodic, sustained)

- c) Professionalization (low, medium, high)
- d) Policy Influence (absent, informal, formalized)

More importantly, it was concluded that VOPEs' contribution to the promotion of a national evaluative culture partly depended on the specific position that they occupied within their respective national "ECD ecologies," as shown by the application of the SFAR Framework (Chapter IV) in the three countries where the case studies were developed. ReNSE, for instance, was the one that appeared to be the most well placed to influence policy and foster dialogue among actors operating in the governmental and non-governmental spheres (true sphere-crossing entity) (Appendix L). SAMEA's role seemed a little bit more limited as it played a very cementing role within its government and non-governmental members separately but it did not seem to be able to enhance an uncomplicated exchange between the two spheres (sphere-border approaching entity) (Appendix M). To the contrary, ACoSE, although it was characterized by a variety of thematic groups and it had established a relatively close relationship with the national civil society platform, did not seem to have played a critical role in the promotion of an evaluative culture as of yet (sphere-bounded entity) (Appendix K), mostly due to its low diversity in both membership spheres (larger share of governmental actors) and functions (consultants working in the PRSP Planning Process within the Ministry of Planning accounted for the majority of members).

2.3. What are the factors characterizing the success or the failure of VOPEs that should be taken into account in view of their involvement in ECD programs in international development contexts in the future?

Based on the three case studies presented in this study, the factors that appeared to most enhance a VOPE's contribution to the development of an evaluative culture, especially at the national level, included the following:

- a) Membership diversity: the more diverse VOPE members in terms of roles and functions were, the more likely it was that a VOPE would affect the national evaluation discourse within the two ECD spheres;
- b) Leadership type: the more decentralized leadership was and the larger the availability of channels through which members could contribute to VOPE decision-making processes was, the better the compliance with internal governance rules was;
- c) Frequency of VOPE capabilities as well as organizational processes: the more frequent—e.g., every year—capability assessments were, and the more promptly the identified weaknesses were addressed, the more successfully VOPE were able to fulfill their missions;
- d) Availability and continued monitoring of a VOPE theory of change: the more often a VOPE Theory of Change was available—e.g., based on the VOPE history and a capability assessment—the more often the assumptions underlying it were monitored, and programmatic improvements were made;
- e) Peer-learning and Peer-exchange: the more frequently exchanges and visits took place both among VOPE membership and between them and those of other institutions (both at the national, regional and international levels), the stronger the sense of community within VOPE was;



- f) Communication strategy: the more well articulated a communication strategy was, the more social and political legitimacy VOPE could benefit from<sup>180</sup>;
- g) Very good diplomatic skills among VOPE coordinators: the more socially savvy VOPE coordinators were, the stronger the VOPE was in taking actions and reaching out to members. The ideal coordinators' skills identified by respondents included the following: high-level professionalism recognized by the national and international community; excellent communication skills (also through the use of social media) diplomatic skills; open-mindedness; independence (no conflict of interest); time availability (he/she does not need a letter of authorization from his director to participate in a meeting/a conference); very good management skills; ethical rigor; and good understanding of equity.

To the contrary, those factors that seemed to have hindered the successful involvement of VOPEs in ECD programming and that would need be adequately mitigated or preventively addressed in the future included:

- a) The lack of adequately targeted initiatives among VOPE members: in order to enhance members' sustained involvement in VOPE activities, the creation of thematic groups combined with the regular organization of General Assembly meetings and VOPE-broad exchanges was of utmost importance;
- b) The low level of government's trust in civil society and the risk of fragmentation or low political and social legitimacy of VOPEs: governmental officers were not eager

---

<sup>180</sup> A key tool enhancing the promotion of a national evaluation culture both within the members and outside of the VOPE (where people might still not have a good understanding of what evaluation is), a good communication strategy was particularly instrumental in attaining two primary objectives. First, projecting the VOPE image nationwide. Second, advocating to some elected officials in the Parliament and get their buy-in during ECD design and implementation.

to collaborate with non-governmental entities actively engaged in attacking and criticizing national policies and departments' work. Actors in the governmental sphere seemed more available to deal with those within civil society who possessed the most advanced technical knowledge of evaluation (e.g., researchers and evaluators);

- c) The lack of an effective decentralization of the M&E function: it was really difficult to promote a defused culture of evaluation through a top-down approach. Through the creation of VOPE chapters at the sub-national level, for instance, some important steps were made to foster evaluation culture further;
- d) The relatively weak data analysis skills among national evaluators: despite the rise in the number of evaluations being conducted in the field, the low capacity in statistical or qualitative analysis seemed to have compromised the quality and credibility of evaluation data, thus discouraging their use;
- e) The widespread belief that evaluation was about verifying the compliance with the envisaged objective with almost no reference to unexpected impact (RBM-ization<sup>181</sup> of the evaluation function): due to such a “reductive” understanding of the evaluation, evaluation risked becoming such a highly centralized and internal function that little room would be left for conducting independent evaluations of programs, especially of the unintended positive or negative impacts of public-funded programs.

---

<sup>181</sup> Results-Based Management.

Other factors hindering the success of ECD identified in the course of the study were:

- The lack of government's dissemination of evaluation findings;
- Low motivation to participate in Association's activities;
- Lack of leadership's time availability

#### Current ECD Frameworks' Limitations and Opportunities for More Inclusive Targeting in the Future

In response to the prominent role played by national governments in both ECD targeting and funding up until the mid-2000s, this study was conducted to identify some opportunities for making ECD targeting more inclusive and effective in the future. In doing so, the limitations of three of the main ECD theoretical frameworks in use were identified and a new framework (SFAR Framework) was developed to fill the observed gaps and enhance in-country ECD in the future. In particular, the SFAR Framework was based on three main principles: a) ECD stakeholders could be both providers and commissioners of evaluation services; b) VOPEs were not a simple ECD strategy but they represent key ECD stakeholders; and c) the three ECD levels (individual, organizational and institutional) were characterized by different but complementary processes; individual-level outcomes could be pursued at each of the three levels.

Overall, it was concluded that, far from being considered as vulnerable recipients of ECD services, government officials as well as representatives of both the civil society and the private sector needed to be increasingly regarded as ECD active consumers and providers. Likewise, actors in both ECD spheres were to be seen not only as committed and

resilient agents of change but also customers who could either buy into or reject the ECD programs or services offered to them.

Deeply rooted in social justice theory (House et al., 1999), this study was aimed at advocating to ECD planners and funders to enhance the opportunities for VOPEs to “sit at the same table” with donors to plan ECD interventions in the future. Similarly, the following was advocated:

- a) The development and adoption of a community-led definition of ECD needs;
- b) The formulation of a shared ECD mission focus among donors/national governments and local evaluation associations;
- c) The establishment of a more responsive governance and administrative structure (both among donors and ECD programs users);
- d) A more active community participation in the definition of development programs;
- e) A more context-specific definition and valuing of capacities; and
- f) A larger community control of evaluation resources.

#### Good Practices to Enhance ECD Inclusiveness

The identification of a new basis for dialogue between donors and development partners (national government and civil society) was one of the objectives envisaged by this study. However, far from either serving as a quick fix to the ECD dilemma or putting forward a cost-effective framework to conduct and evaluate ECD, this study was intended to

provide ECD programmers with a more operational ECD framework<sup>182</sup>. As a result, a roadmap for a successful ECD programming and evaluation was developed (Box 1). Rather than being descriptive (as in the case of a recipe book), the suggestions included in this list were intended to be adapted to the specific context where ECD would be implemented.

**BOX 1. A Few Useful Steps to Put in Place for a More Inclusive and Successful ECD**

**1) Understand the specific ECD ecology where you are working.** In conducting a mapping of the major institutions (both within the government and non-government sphere, including VOPE, academia, private sector) that have demonstrated interest in ECD in the past – the identification of individuals as well as specific units is strongly recommended.

**2) Identify some common nationwide goals and objectives for all ECD stakeholders to contribute to.** Such goals, better if aligned with Millennium Development Goals (e.g., MDG until 2015), do not need to be perceived as imposed from the outside and should be consistent with the in-country actors' mission and objectives<sup>183</sup>.

**3) Build institutional incentives from within.** The conduct and use of evaluation findings (with the corresponding budget allocations) are allocated in Sector-Wide Approaches, Pooled Funds, PRSP, ECD country-level agreements to which national governments are signatories and asked to provide counterpart funding which would result in the ownership and sustainability (not only financial but also programmatic and institutional and adaptive stewardship) being greatly improved. The incentive made available, including the identification and rewarding of champions<sup>184</sup>, should not focus

---

<sup>182</sup> To quote Baddoo, Ubels and Fowler: "Further dialogue is required to bring coherence (not homogeneity), structure (not control) and rigor (not prescription) into the way capacity development practitioners, their organizations and funding partners act and think together as co-shapers of this area of work" (Badoo et al, 2010, p. 45).

<sup>183</sup> This is in line with a popular OECD publication released in the aftermath of the Paris Declaration and calling upon ECD programmers to: Harmonize their monitoring and reporting requirements, and, until they can rely more extensively on partner countries' statistical, M&E systems work with partner countries to the maximum extent possible on joint formats for periodic reporting. (OECD, 2005, p.34)

<sup>184</sup> The World Bank-funded Capacity Building Project in the DRC seems to have recognized the value of champions' individual commitment towards the successful implementation of a nationwide ECD program. To this end, the US\$30 million project has envisaged specific leadership programs targeting capacity development champions within a number of ministries. Besides providing them with motivational training, the program is also providing them with coaching address their respective organizations' capacity needs (based on the findings of a diagnostic conducted in 2010 and 2011). Overall, the project has four main components: (i) strengthening leadership and introducing basic results-based management tools; (ii) strengthening core public administration functions; (iii) support to local training centers to define and implement programs of excellence, and; (iv)

on meeting pre-set performance agreement and should promote instead the ECD actors' ability to wonder and ask questions on how to turn tactics into strategies.

**4) Conduct a participatory ECD diagnostic assessment.** The questionnaire used to identify the evaluative capacity of VOPEs in the three countries where the three studies presented in Chapter IV were conducted (DRC, Niger and South Africa) might be useful. Through such assessment, it will be important to foster opportunities for self-reflection and mutual learning.

**5) Develop a national ECD strategy and, depending on the scope of your ECD intervention, put in place a national ECD Task Force.** Far from being a logical framework or road map, a national evaluation strategy is to be regarded as a work in progress and a living document, setting general objectives and leaving ECD stakeholders space for them to come up with a creative and innovative way to achieve the agreed upon objectives. This phase might require the address of some key issues such as the creation of dedicated evaluation units with three specific responsibilities (compiling a database of evaluation data, conducting data analysis and, foster dissemination of evaluation findings) as well as partnerships between departments. With respect to the ECD Task Force, it is advisable that a variety of actors with different functions and roles (from both spheres) as well as sphere-crossing entities, be involved in it.

**6) Focus on strengthening the capacity of local actors (both users and providers of evaluation services) and develop an opportunistic joint exit strategy, in close collaboration with the individuals, organizations and institutions involved in a ECD program.** Although not operationalized from the outset, an exit strategy needs to be in place. In order to advance sustainability, it is of utmost importance to have a very strong leader in place, who is capable of innovation and available to promote internal structure changes and the reallocation of budget resources based on the organization's needs and the changing contextual opportunities (referred to in this study as "responsible systemic-ness").

**7) Evaluate the progress of your ECD over time.** In order to measure the effectiveness of your ECD support program, you need to make sure that your target audience as well as the type of capacity whose development you are trying to support (based on the interests and needs expressed by in-country ECD stakeholders) have been clearly defined at the time of ECD strategy development.

### Third Research Question

How Can, or How Should, ECD Best Be Evaluated?

3.1 What are the key criteria that need to be taken into account in order to assess ECD effectiveness?

Based on both a systematic literature review and a series of semi-structured interviews conducted in the field with a variety of evaluation practitioners and commissioners, four sets of variables commonly used to measure ECD effectiveness in the past were identified.

The first set of variables, the most commonly used of the four presented in this section, was aimed at assessing the short-term results (such as, increase in knowledge and development of technical skills) produced by ECD activities (e.g., training, coaching, mentoring) implemented at the individual level.

The second set of variables was aimed at assessing the medium- and longer-term results produced by ECD activities implemented at the organizational and institutional levels (e.g., the adoption of mid-course corrections based on evaluation findings<sup>185</sup> or the Parliament's approval of a national evaluation policy).

The third set of variables was aimed at assessing the internal processes inherent to evaluation providers<sup>186</sup> and evaluation consumers<sup>187</sup> (e.g., type and quality of interactions between them, degree to which the same targeted individuals could play both roles at once).

---

<sup>185</sup> Henry and Mark speak of three different type of utilization: a) Individual (attitudinal change, skill acquisition, and behavioral change; b) Interpersonal (changes in justification, persuasion and social norms), c) Collective (changes in setting the agenda and modifying policy) (Henry & Mark, 2003).

<sup>186</sup> Evaluation of processes inherent to ECD providers focused, among others, on the degree of ECD targeting's inclusiveness, the level of budget resources available for implementation of activities in the field or the timeframe within which ECB activities are expected to yield their envisaged results.

<sup>187</sup> Evaluation of processes inherent to ECD consumers focused, among others, on special traits of the population whose capacity was being developed, including their needs, interests and motivation level.

The fourth set of variables was aimed at assessing the relevance of contextual factors (e.g., the institutional arrangements in place within organizations being capacitated or the special political system where the ECD program was taking place) and their mediating effects on ECD outcomes.

Overall, it was concluded that, rather than assessing the change in specific outcomes related to the actual delivery of training and mentoring, what appeared more important was to look at the change in those external factors that were identified as most affecting changes in capacity. In addition to the more general criteria used to evaluate the effectiveness of comprehensive ECD programming, some more specific variables were identified to assess VOPEs' capacity and their effectiveness on the promotion of an evaluative culture. Based on the retrospective formulation of a VOPE theory of change and the identification of five main outcomes associated with their activities, a corresponding set of evaluative criteria were proposed to measure the extent to which VOPEs were able to attain their primary objectives (Table 17) . Variables under this category included the following:

- a) Knowledge Sharing (discontinued, episodic, systematic)
- b) Knowledge Building (absent, episodic, sustained)
- c) Professionalization (low, medium, high)
- d) Policy Influence (absent, informal, formalized)
- e) Strengthening of Institutional Relations (low, medium, high)



Table 17

*VOPEs Main Objectives and Outcomes: A Comparative Summary*

Country	DRC	Niger	South Africa
Criteria			
Knowledge Sharing	*	**	***
Knowledge Building	*	*	**
Research			
Professionalization	*	*	***
Policy Influence	**	***	**
Defend the interest of evaluators as professionals networks	*	**	***

## The Relevance of Process-related Outcomes

Based on the interviews with in-country evaluation practitioners and the understanding that processes are key to the success of ECD programs, a list of evaluation questions, radically different from the more compliance-related ones currently in use, were identified and recommended to be adopted in the future:

1. To what extent did the ECD intervention planners carefully analyze the motivations, transformational moments, fears, wants and needs of the targeted actors?
2. To what extent did ECD planners 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?
3. To what extent were ECD planners able to identify an untapped, willing momentum for change?
4. To what extent did ECD progresses, delays, insights, assumptions and activities the object of participatory discussions?

5. To what extent did the design of various capacity strengthening activities vary often enough to keep responding to emerging needs?
6. To what extent did the ECD planners experiment enough?
7. What is known now and that was not known at the beginning of the ECD intervention?
8. To what extent are beneficiaries involved in the evaluation process?
9. To what extent did ECD rely on local evaluators or, when not possible, on third party evaluators from partner countries?
10. To what extent does the funder engage in partner-led evaluations?
11. How did partners/beneficiaries/local NGOs perceive the evaluation processes and products promoted by the agency/country examined in terms of:
  - quality,
  - credibility
  - independence,
  - usefulness and partnership orientation?

-

#### Five Key Principles in Evaluating ECD Programming

Based on the data analysis, five ECD guiding principles were identified:

1. Evaluation of ECD processes (rather than performance only) is critical: assessing the type and quality of processes inherent to ECD stakeholders (e.g. inter-organizational dynamics, availability of “learning space” within an institution, VOPEs resilience at

time of crisis) is key to identifying a priori factors that may either enable or hinder ECD;

2. Assessing the level of ECD needs, interests and motivation across the three levels is critical to customizing and sequencing ECD programs;
3. Gauging the quality and the degree of ECD targeting's inclusiveness is instrumental in the identification of any possible inequity (and strategy available to address it) in ECD programming;
4. Dissemination of details on the budget resources available for and effectively spent for the implementation of ECD programming is key to cost-benefit and cost-effectiveness analyses;
5. Assessing the feasibility of ECD success given the established timeframe is relevant

#### Sequencing of ECD Evaluation

Evaluations of ECD programs are expected to be conducted at multiple stages during the implementation process. Four main phases were identified in the course of this study a series of specific actions for each of them were highlighted:

##### Before the implementation of an ECD program:

1. Conduct a multi-level capacity needs assessment;
2. Conduct a diagnostic of processes at each of the three level (use secondary data as appropriate); and
3. Gauge the level of Monitoring knowledge and familiarity with Results-based management (RBM) (Jalijouli, 2011)

##### During of right after implementation:

1. Assess short-term results (such as, increase in knowledge and development of technical skills) produced by ECD activities for each of the three levels.

Shortly after implementation:

1. Start assessing systematically medium- and longer-term results produced by ECD activities implemented at the organizational and institutional levels;
2. Conduct individual or group follow-up interviews and make the best use of online tools to foster reflections and conversations on the lessons/challenges resulting from ECD programs

Through all ECD programming phases:

1. Assessing the internal processes (e.g., type and quality of interactions among actors in different spheres and with different functions as well as the degree to which the targeted individuals could be both providers and consumers at once);
2. Monitor the assumption underlying your Theory of Change;
3. Be very systematic in your assessment of the mediating factors (both those included in your theory of change and others that you might identify during implementation)

### 3.2 At what level do the identified key ECD criteria need to be evaluated?

Based on the field interviews with a variety of ECD stakeholders as well as content analysis of the ECD specialized articles included in the systematic literature review, it became apparent that, for evaluating ECD effectiveness adequately, it was not sufficient to measure a list of pre-identified evaluative variables measuring effects at the individual level (e.g., level of knowledge increase or level of adoption of good evaluation practices). Rather, it was of utmost importance to measure the changes occurring both at the organizational (e.g., including the degree to which incentives and knowledge management processes were able to

enhance aimed more consistent practice and use of evaluation within the organization) and institutional levels (e.g., the degree to which evaluation is integrated across both spheres often driven by a national evaluation policy and a national capacity development strategy).

On a more general note, it was concluded that assessing ECD effectiveness was not only about levels but also, and more importantly, about sub-levels. Otherwise said, it became apparent that it was critical to be able to identify the specific unit of analysis within each of the three levels at which the outcome of interest needed to be measured (e.g., level of knowledge increase among individual commissioners of evaluations disaggregated by ministry and sex or level of evidence-informed planning across an entire organizations).

Consistent with the SFAR Framework predicament, it also became apparent that an example of a good ECD indicator would be the number and percentages of individuals targeted by ECD programs, disaggregated by sphere (governmental/non-governmental) as well as their respective type of functions (operational, policy-making, decision-making) and roles (consumers, provider, initiator/commissioner). As a result, for each of the four sets of evaluative variables presented in this section, the level at which ECD outcome of interest needed to be measured was clearly indicated.

### Study Limitations

Four main limitations were identified in relation to this study. First, although a special effort was made to gather the ideas and opinions on ECB and ECD among as many in-country evaluation practitioners and scholars as possible, so as to yield generalizable (Schofield, 2002) or transferable findings on ECD across a variety of contexts<sup>188</sup>, the focus

---

<sup>188</sup> Some might link this effort to the methodology know as “analytic” or “case-to-case generalization” (Firestone 1993, p. 18) or “petite generalization” (Creswell, 1998; Erickson 1986, p.120)

of this study remained on three specific countries (DRC, Niger and South Africa). Therefore, despite the fact that the critical or extreme features characterizing the VOPEs discussed in this study might be found in VOPEs with the same level of capacity or operating in the presence of similar (but not identical) contextual factors, the applicability of the conclusions associated with each of the case studies might be limited. Thus, caution is needed in extrapolating the results of this study to other contexts, or in drawing general conclusions about the ECB-related factors and relationships.

Second, although the SFAR Framework was applied to the three different case studies included in this study (DRC, Niger and South Africa), the validity of the framework could only be enhanced by testing it (and revising accordingly) to a larger number of countries. The fact that all three countries on which the SFAR Framework was tested are in sub-Saharan Africa also represented a main limitation in that some of the SFAR principles would not be able to fully capture the reality of VOPEs operating in other contexts (e.g., Latin America, Asia or the former Soviet Union).

Third, by contacting VOPEs through the database posted on the IOCE website and by establishing a link with the coordinator and the members that he or she recommended (although the study was advertised widely in Niger and DRC before the start of data collection), the ideas collected among VOPE members in the course of data collection may be representative of a restricted group of VOPE members and not of the whole membership.

Fourth, although the main strengths and weaknesses of VOPEs in three different countries were described in this study and an effort was made to build on such information to provide a better explanation of the way VOPEs function, this remained an exploratory

study. As a result, despite the suggestion that VOPEs ought to be more involved in ECD programming in the future, no relationship was established between the degree of VOPE involvement and the state of evaluation culture in any of the three countries where the case studies were conducted.

### Implications for Theory and Practice

Overall, this dissertation represented an ambitious and yet humble effort to shift the vision of ECD in the international development discourse from a predominantly functional or compliance-drive perspective to a more democratic and systemic paradigm, by also suggesting new questions that practitioners and scholars in the field might want to address in the future.

As Morgan would say (Morgan, 2006, p. 7): “Most practitioners are in favor of capacity as an idea. But only a few among them actually talk in specific and strategic ways.” Well embedded within the broader ECD discourse, this study was aimed at bringing about some innovative ideas, thanks also to the validation process involving several representatives from international organizations and national evaluation associations.

If it is true that a key bottleneck to the success of ECD is not so much the lack of capacity but a weak evaluation culture (Segone, 1998), this study represented an attempt to address the so-called contextual and environmental factors more systematically in the analysis of VOPEs<sup>189</sup>. More in particular, in an effort to address the limitations of the existing definitions (e.g., the lack of operational definitions of key ECD concepts), a more easily understandable and practical framework for action in the ECD arena was provided.

---

<sup>189</sup> Picciotto (1998, 2011) echoes Segone’s words in asserting that the pivotal role of context, more precisely the lack of an evaluation culture, as a strong deterrent to ECD. However, he – as well as other authors – did not seem to make a strong case for ECD to claim a much broader scope, including addressing issues associated with the enabling or hindering environment.

While the variety of insights emerged from this study were not intended to be of direct applicability to other contexts than those of the three countries where data collection was undertaken, they were likely to help researchers and practitioners to better understand the nature of ECD, as well as to inform future studies and work in this field. The results of this study were aimed at contributing to the body of ECD research by focusing on ECD process factors and the relevance of implementing holistic interventions.

Consistent with the study objectives, the findings of this work are expected to be disseminated widely. A few weeks before the defense, the preliminary results of this work had been presented for feedback and preliminary validation in several settings, including a series of meetings held with evaluation officers in international aid agencies (e.g., UNDP, OECD and UNICEF) and three presentations at the European Evaluation Associations Biannual conference (Helsinki, October 2012) and American Evaluation Association Annual Conference (Minneapolis, November 2012). The key preliminary findings of this work were also shared back with the coordinators and members of the three VOPE on which the case studies presented in Chapter IV had been developed. The dissemination plan included (a) the publication of the study results in peer-reviewed journals; (b) a number of brown bag presentations at several international development agencies; (c) presentations on ECD to VOPE around the world; and d) presentations during professional conferences in 2013-2014.

#### Future Research

A number of future research topics were identified upon completion of this study. First, the opportunity for conducting some cross-analysis between the findings of the case studies presented in this study and other VOPE case studies being conducted either



concurrently or after the completion of this study (e.g., those developed by EvalPartners and due to publication in late 2012 or those expected to be developed by the Geneva-based Evaluation Capacity Development Group in 2013).

Second, based on the findings of this study, some further studies to measure the effectiveness of different ECD programs characterized by different degrees of inclusiveness could be conducted. In particular, it would be beneficial to compare the effects of programs targeting individuals with different roles and functions in both spheres (as predicated by the SFAR framework) with two other types of interventions: one intentionally targeting only one sphere and the other one targeting actors in both spheres but either at the operational or decision-making level.

Third, the influence that RBM trainings and MfDR initiatives have on the general understanding and implementation of the evaluation function both in sub-Saharan Africa and other regions of the world (where CoP-MfDR exist) is a topic that might be worth exploring further in the future.

Fourth, the issue of capacity could be researched further in the future by making sure to use frameworks borrowed from other disciplines. Such is the case of (a) frameworks currently used in business management and potentially useful to explore all the questions related to leadership and VOPE organizational development; and (b) frameworks used in political economy that could be used in the future to inquire on the feasibility and effectiveness of incentives and the promotion of an evaluative culture at the institutional level.

Fifth, social network analysis could also be adopted in the future to better map and understand the frequency, direction and magnitude of interactions existing between VOPEs and the rest of the ECD actors in the three countries where the case studies were conducted.

## Appendix A

### Acronyms

## Acronyms

This appendix is intended to serve as a point of reference for the many acronyms that occur throughout the dissertation

ACoSE: Congolese M&E Association  
AEA American Evaluation Association  
ACBF African Capacity Building Foundation  
AfCoP African Community of Practice  
AfrEA African Evaluation Association  
AusAID Australian Agency for International Development Agency  
AJE American Journal of Evaluation  
ADB Asian Development Bank  
BMZ German Federal Ministry of Economic Development Cooperation  
CD Capacity Development  
CDRA Community and Development Organisation for Social Change  
CES Canadian Evaluation Society  
CIDA Canadian International Development Agency  
CoP Community of Practice  
COSATU Congress of the SA Trade Unions  
CPF Country Programme Framework  
CREST Center for Research, Evaluation, Science and Technology  
CSO Civil Society Organization  
CTA Complex Adaptive Thinking  
DBE Department of Basic Education  
DPE Planning and Research Unit  
DFID UK Department for International Development  
DPME Department of Performance Monitoring and Evaluation  
DRC Democratic Republic of Congo  
ECB Evaluation Capacity Building  
ECD Evaluation Capacity Development  
ECDPM European Center For Development Policy Management  
ECoP Evaluation Community of Practice  
ECOSOC United Nations Economic and Social Council  
EES European Evaluation Society  
ENACT Environmental Action Program in Jamaica  
FAO Food and Agriculture Organization  
FMED Fragile Market Economies Division  
GIZ German International Development Agency  
HR Human Resources  
HSIRB Human Subjects Institutional Review Boards  
IDEAS International Development Evaluation Association  
IDPE Interdisciplinary Doctoral Program in Evaluation

IDRC International Development Research Center  
IDS Institute of Development Studies  
IEC The Independent Electoral Commission  
IEG Independent Evaluation Group  
IFAD International Fund for Agriculture Development  
IOCE International Organization for Cooperation in Evaluation  
ISEP Higher Institute of Evaluation and Perspectives  
IUCN International Union for the Conservation of Nature  
LECDP Local Evaluation Capacity Developer  
MDG Millennium Development Goals  
M&E Monitoring and Evaluation  
M&ES Monitoring and Evaluation Staff  
MfDR Management for Development Results  
MPAT Management Performance Assessment Tool  
NEPF National Evaluation Policy Framework  
NGO Non-Governmental Organizations  
NORAD Norwegian Agency for Development Cooperation  
OECD Organization of Economic Co-operation and Development  
OA Outcome Approach  
OCBD Office of Capacity Building and Development  
ODA Overseas Development Assistance  
PALAMA Public Administration Leadership and Management Academy  
PPL/LER USAID Office of Learning Evaluation and Research  
PPP Processes, Policies and Practices  
PSC Public Service Commission  
PRSP Poverty Reduction Strategy Paper  
RBM Results-Based Management  
ReNSE Niger Monitoring and Evaluation Network  
ROM Results Oriented Management  
SA South Africa  
SAENet South Africa Evaluation Network  
SAHRC The South African Human Rights Commission  
SAMEA South African Monitoring and Evaluation Association  
SASSA South African Social Services Agency  
SENAREC National Secretariat for Capacity Building  
SFAR Sphere-Function-Actors-Role Framework  
SFE French Evaluation Society  
SIDA Swedish International Development Agency  
TA Technical Assistance  
TAU Technical Assistance Unit  
TIG Topical Interest Group  
ToR Terms of Reference  
VOPE Voluntary Organization of Professionals in Evaluation  
UCT University of Cape Town  
UN United Nations  
UNDP United Nations Development Programme  
UNEG United Nations Evaluation Group

UNICEF United Nations Children's Fund  
UNIDO United Nations Industrial Development Organization  
UNESCO: United National Educational, Scientific and Cultural Organization (UNESCO)  
Correct on page 33 from National to United  
USAID United States Agency for International Development  
USDA United States Department of Agriculture  
WB World Bank  
WBI World Bank Institute  
XEval International and Cross-Cultural Evaluation List serve

## Appendix B

### Semi-structured Interview Protocol

## QUESTIONS TO SELECTED ECD SPECIALISTS, SCHOLARS AND VOPEs

1. How do you define evaluation capacity building (ECB)?
2. Could you provide an example of ECB in your country?
3. Do you ever use the term Evaluation Capacity Development (ECD)?
4. How do you see the relationship between ECB and evaluation capacity development (ECD)?
5. What are some of the most important challenges that donors will need to address during the implementation of new ECB and ECD initiatives and programs in the future?
6. What are some possible cost-effective solutions to address some of the challenges listed above and enhance the effectiveness of ECD endeavors in the future?
7. What are the three main reasons you joined a VOPE?
8. What type of activities does the national evaluation network get involved in?
9. How could VOPE get involved more effectively in the promotion of an evaluative culture in their respective countries?
10. What are some contextual factors enhancing the effectiveness of ECB/ECD programs that target in-country partners' national evaluation associations/networks in your country?
11. What are some contextual factors hindering the effectiveness of ECD programs that target in-country partners national evaluation associations/networks in your country?
12. What are the key ECD dimensions that need to be taken into account in order to assess ECD effectiveness?
13. What are the new topics that future ECB/ ECD endeavors should include in current trends in evaluation practices among members of national evaluation associations/networks (e.g., most popular approaches, level of understanding and implementation of different research designs and methods)?
14. How are capacity building activities evaluated in your country?

## Appendix C

### The Five Key Dimensions of the ECPDM Capability Assessment



## The Five Key Dimensions of the ECPDM Capability Assessment

The first capability (Capability to Commit and Act) is the ability of an organization to project its identity, mission and values both internally (among its members) and externally (ECD stakeholders within the government, the donors' community as well as the general population). This is the so-called latent capacity, that is, the foundation of an organization that is the necessary condition for all other types of abilities to develop. Also it can be referred to as the ability of an organization to function as a resilient<sup>190</sup>, strategic and autonomous entity (Kaplan, 1999).

The second capability (Capability to Generate Development Results), is also referred to as the function, technical and logistical ability to get work done (Morgan, 2006) and is often understood as the ability to enhance on an organization's performance, as attested by some tangible outputs and outcomes, especially if related to service delivery. This is interpreted as the core antidote to local capacity gaps and deficiencies (business-like approach), any activity or program enhancing this capacity is more reactive in nature and overlooks the causes of such deficiencies. The great focus on products in short-term rather than medium- and long-term processes discourages loop learning or introspective and reflective exercises leading to organizational changes;

The third capability (Capability to Relate) refers to organizations' ability to interact pacifically and forge alliances with other actors within the system where they operate, so as to create a buffering zone where they could feel protected and, thanks to a well-recognized legitimacy within society, gather sufficient support for the continuation of their activities;

The fourth capability (Capability to React) is critical in an increasingly chaotic world and is a necessary attribute for organizations that desire a means to effectively address uncertainties. The ability to react consists in adapting and modifying plans and operations based on monitoring of progress and outcome as well as in proactively anticipating change and new challenges. This capacity also concerns the ability to cope with shocks, develop resiliency, foster internal dialogue and incorporate new ideas.

---

<sup>190</sup> Resilience is gradually replacing the old paradigm of vulnerability in international development. This is no exception in the UN, one whose agencies (UNDP) recently developed a working definition of the term resilience: "a transformative process which builds on the innate strength of individuals, their communities, and institutions to prevent, lessen the impacts of, and learn from the experience of shocks of any type, internal or external, natural or man-made; economic, health-related, political or social. The definition also outlines the five key operating principles of resilience: respect for context-specificity and national ownership, comprehensiveness and integration, partnership, commitment to innovation and learning, and strategic and long-term engagement".

The fifth capability (Capability to Integrate): the ability of an organization to fully exploit the opportunities available outside of the organization without compromising its own principles and value. Signs attesting to such capacity are the reliance on “cross-functional, cross-country, cross-disciplinary team and management groups” (Fowler & Ubels, 2010, p. 21)

Appendix D

VOPE Capability Assessment Tool

VOPE Capability Assessment Tool:

<p><b>1) Capacity to Commit and Act (latent capacity): VOPE demonstrate collective energy and mobilize others to act</b></p>
<ul style="list-style-type: none"> <li>- Does your association have a mission?</li> <li>- Does your associations have values agree upon and recognized as their own by all its member?</li> <li>- Does your association dispose of a Work Plan? If yes, how did you develop and who got involved?</li> <li>- To what extent is your association adopting collective decision-making practices?</li> <li>- How does your association mobilize resources?</li> <li>- To what extent is your association monitoring its Action Plan?</li> <li>- What is the type of leadership (inspiring/action oriented), which your association is inspired to?</li> <li>- What is your association’s level of confidence about being successful in the pursuit of its mission?</li> <li>- To what extent does your association feel autonomous?</li> </ul>
<p><b>2) Capacity to Generate Development Result (functional, technical and logistical capacity): VOPE attain outputs and outcomes and sustain delivery by also adding value for their members.</b></p>
<ul style="list-style-type: none"> <li>- What is the level of your financial Resources within your association?</li> <li>- What is the level of human Resources within your association?</li> <li>- What is the level of access to Knowledge resources within your association?</li> <li>- To what extent are your association’s activities aligned with your country’s National Development Strategy? PRSP (Indicator 4 Paris Declaration)</li> </ul>
<p><b>3) Capacity to Relate: VOPE can forge alliances and partnerships with others to leverage resources and action; build legitimacy in the eyes of key stakeholders; and deal effectively with commotion, politics and power differentials</b></p>
<ul style="list-style-type: none"> <li>- To what extent does your association participate in coalitions?</li> <li>- To what extent does your association forge alliances with external stakeholders?</li> <li>- What is your organization’s level of social legitimacy?</li> <li>- What is your organization’s level political legitimacy?</li> </ul>
<p><b>4) Capacity to Adapt: VOPE are able to adapt and modify plans and operation based on monitoring of progress and outcomes; proactively anticipate change and new challenges; cope with shocks and develop resiliency; and foster internal dialogue and incorporate new ideas.</b></p>
<ul style="list-style-type: none"> <li>- To what extent does your association understand shifting contexts and relevant trends?</li> <li>- To what extent did your association modify its plans or operational plans in response to a contextual analysis?</li> </ul>

- To what extent does your association demonstrate confidence to change (e.g., leaving room for diversity, flexibility and creativity)?
- To what extent does your association use opportunities and incentives and acknowledge mistakes that have been made to learn?
- To what extent does your association systematically plan and evaluate learning including management?

**5) Capacity to integrate: VOPE can develop shared short and long term strategies and visions; balance control, flexibility and consistency; integrate and harmonize plans and actions in complex multi-actor settings; and cope with cycles of stability, change and innovation**

- To what extent is your association's leadership committed to achieving coherence between openness and operating principles?
- To what extent does your association balance stability and change?
- To what extent does your association demonstrate coherence between ambition, vision, strategy and operations?

*Source:* Adapted from Morgan (2006) and European Center For Development Policy Management Core Capabilities (2011)

Appendix E

HSIRB Approval Letter

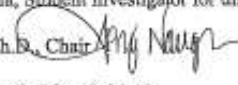
# WESTERN MICHIGAN UNIVERSITY



Human Subjects Institutional Review Board

Date: April 12, 2012

To: Chris Coryn, Principal Investigator  
Michele Tarsilla, Student Investigator for dissertation

From: Amy Naagle, Ph.D., Chair 

Re: Approval not needed for 12-04-13

This letter will serve as confirmation that your project "Aid Effectiveness from Words to Action\*: How to Enhance National Ownership of Evaluation Capacity Development (Ecd) Programs in Developing Countries: The Case Studies of South Africa, The Democratic Republic Of Congo And Niger" has been reviewed by the Human Subjects Institutional Review Board (HSIRB). Based on that review, the HSIRB has determined that approval is not required for you to conduct this project because you are analyzing evaluation capacity development programs and not conducting human subject research.

Thank you for your concerns about protecting the rights and welfare of human subjects.

A copy of your protocol and a copy of this letter will be maintained in the HSIRB files.

Walwood Hall, Kalamazoo, MI 49008-5456  
PHONE: (269) 387-8293 FAX: (269) 387-8276

Appendix F

ECD Continuum Feedback Questionnaire



## ECD Continuum Feedback Questionnaire

### **ECD Continuum Feedback Questionnaire**

Please type in your feedback below and return this form to [michele.tarsilla@wmich.edu](mailto:michele.tarsilla@wmich.edu). Use as much space as you need for your answers. Please keep in mind that the intended users of the ECD Continuum are national stakeholders working within or outside the government as well as development partners and VOPE involved in the ECD programming. Thank you.

1. Are the categories of the Continuum sufficient? Should some categories be deleted? Should two or more categories be merged?
2. Should categories be merged? Which ones?
3. Does scoring each category from 1 to 5 (1 = very limited consideration of the item; 5 = very high consideration of the item) make sense? How could it be improved?
4. What else should be changed about the Continuum?

Thank you very much for your feedback!

## Appendix G

### Review of Literature on Capacity: Summary Key Findings

## Review of literature on Capacity: Summary Key Findings

Capacity Definitions	Source
- Instrumental Perspective	Goodman (1998); LaFond and Brown (2003)
- Capacity belongs to individuals, organizations and societies - Needed for sustainable results - Key to problem solving	UNDP (2002)
- Latent, technical, participatory nature of capacity - Accountability purpose of capacity	(Lusthaus et al, 2003).
- Latent nature of capacity	Horton & al., 2003
- Capacity as a prerequisite of Effectiveness and Sustainability	(LaFond and Brown, 2003).
- Holistic concept (people and society as opposed to individuals and societies)	OECD (2006)
- Latency of capacity regardless of its actual use - Uncertainties and unpredictability of the conditions which capacity depends on	<i>Morgan (2006)</i>
- Emerging nature of capacity - Introduction of new related concepts, such as individual competencies and collective capabilities - Systemic (capacity is relational) and value-based perspective in the capacity discourse	<i>Morgan &amp; Baser, (2008)</i>
- Capacity as an internal and critical feature of a human system - Capacity as the basis for survival and self-renewal.	<i>ECDPM (2008)</i>
- Standing nature of capacity. Need for distinguishing between “basic functionalities” “organizational talents” (e.g., permanent and effective qualities and resources)	<i>Ortiz &amp; Taylor, 2009</i>
- More manageable level of analysis (groups of organizations rather than society). - Relevance of context and need for a more thorough analysis of internal dynamic and external processes - Relevance of leadership, loyalty, ambition, power, that is, the relational, psychological and social attributes of capacity.	<i>Boesen, 2010</i>
- “Generic” and “Specific” capacities	<i>OECD 2006</i>
- “Operation” and “adaptive” capacities	<i>Earl, Carden and Smutyllo; 2001</i>

- “Hard” and “Soft” capacities	<i>Horton et al., 2003</i>
- Capacity to: <ul style="list-style-type: none"><li>• Commit and engage,</li><li>• Adapt and self renew,</li><li>• Relate and attract,</li><li>• Balance diversity and coherence</li><li>• Carry out technical, service delivery and logistical tasks</li></ul>	<i>ECDPM, 2010</i>

## Appendix H

### Review of Literature on Capability: Summary Key Findings

Review of literature on Capability: Summary Key Findings

Capability Definitions	Source
<ul style="list-style-type: none"> <li>- Collective nature of the term: capability is uniquely applicable to institutions (not individuals or the society more in general);</li> <li>- Relational feature of any given organization (demand and supply of service);</li> <li>- Emphasis on the relevance of skills and facilities (micro-level variables) as well as legislation (macro-level variable): an environmental scan is needed.</li> </ul>	UNIDO (1990)
<ul style="list-style-type: none"> <li>- Emphasis on the individual (rather than collective) nature of capability;</li> <li>- No link with any specific objective, individual or contextual variable or function.</li> </ul>	Sen (1999)
<ul style="list-style-type: none"> <li>- Renewed emphasis on the collective nature of capability;</li> <li>- No relation to any specific institution but rather to a group or system;</li> <li>- Emphasis on the informal and holistic features of the entities articulating and using capabilities as part of their regular functioning;</li> <li>- Introduction of “generative” capabilities whose intangible features and evolving nature push for more in-depth evaluation conducted over a longer period of time than traditional assessments.</li> </ul>	Morgan (2006)
<ul style="list-style-type: none"> <li>- Distinction between capacity (the latent ability of an individual to do something) and capabilities (the actual ability to use one’s own capacity in practice);</li> <li>- Points to the need for programs aimed not only at capacity (organizational development often taking a longer period of time) but also capability (that is, a set of defined skills of immediate applicability in the interest of the organization in question).</li> </ul>	Bob Williams (2011)
<ul style="list-style-type: none"> <li>- Distinction between capability (regarded as an individual feature) and capacity (regarded as an organizational feature).</li> <li>- Points to the need for acknowledging the dualism inherent to any activity aimed at supporting individual and organizations.</li> </ul>	Ian Davies (2011)
<ul style="list-style-type: none"> <li>- Confirms the distinction between capacity and capabilities</li> <li>- Stresses the need for programs aimed at <i>both</i> greater capacity (e.g., the introduction of new evaluation policies) or greater capabilities (the readiness to commit to action in response to the new policies)</li> </ul>	Heider (2011)

## Appendix I

### Review of Literature on Capacity Development: Summary Key Findings

## Review of literature on Capacity Development: Summary Key Findings

Capacity Development Definitions	Source
- Considers the commonalities between the perspective of capacity offered by it (field implementation and OD are strictly inter-dependent) and the capacity-capability dualism inherent	Kaplan et al (1994)
- Stresses the relevance of know how, that is, the readiness to maximize current and future knowledge and skills to enhance performance at several levels, over the development of a definite set of skills;	CIDA (2000)
- Centrality of ownership - More limited and strategic role of donors - Calls for enhanced customization and adaptability of CD programs to local Needs;	CIDA (2000)
- Definition acknowledges the relevance of both the resources and management available within the organization as well as of contextual factors	Lusthaus et al, (2003)
- Functional or instrumental view of capacity - Not clear distinction between groups and organizations. - The association of changes in governance with only one of the three levels (the institutional level) seems to be a bit too reductive	ADB (2004)
- Ineffectiveness of trainings - The real value of CD as opposed to a label conveniently attached to donors' agenda in order to facilitate the buy-in of host countries	Morgan (2005b)
- It suggests that CD is more of a process than an activity - More recognition of the ecological influences on CD	OECD (2006)
- CD brings about transformations that empower individuals	UNDP (2009)
- CD is the engine of sustainable human development	UNDP (2009)

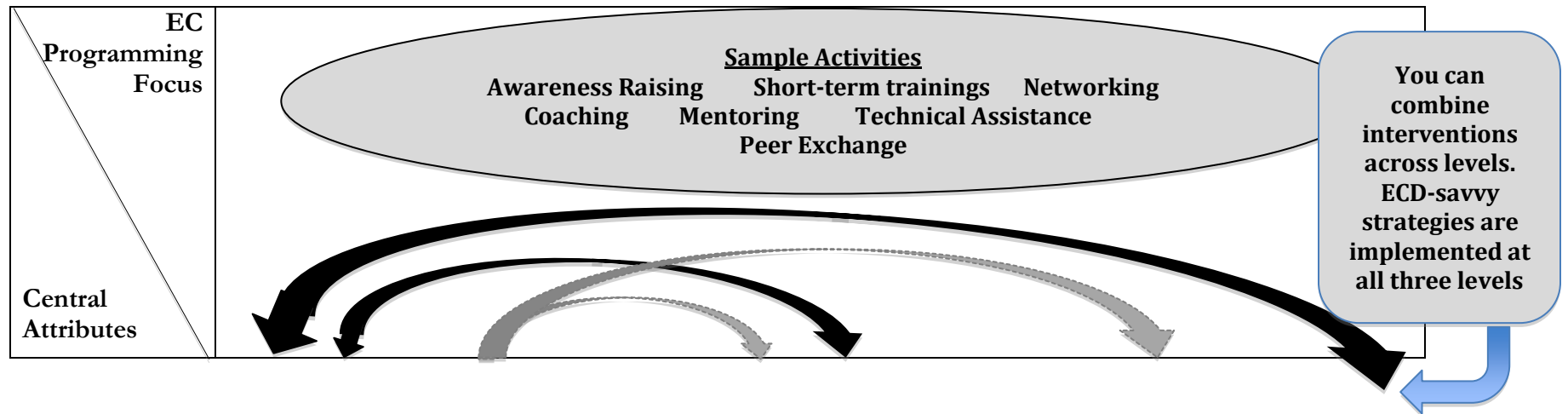


## Appendix J

### The Evaluation Capacity Development Continuum Framework (ECD-CF)

## The Evaluation Capacity Development Continuum Framework (ECD-CF)

(The content of this table has been informed by the literature review readings and interviews conducted in the field.)



<b>Level</b>	<b>Individual</b>	<b>Organizational</b>	<b>Institutional/National</b>
<b>Use</b>	Functional use of short-term training and other punctual actions aimed at increasing evaluation knowledge, skills, and attitudes among individuals. In the past, the primary objective of interventions at this level was to facilitate the smooth implementation of evaluation reporting tasks from field-level staff to funders (donors)	Functional use of capacity strengthening activities geared towards the development of internal organizational capabilities, as well as structures and processes, promoting the evaluation function	Functional use of capacity strengthening activities geared towards the development of norms and policies promoting the planning, management, conduct and use of evaluation by government and other agencies (external to the VOPE itself)
<b>Goal</b>	(Functional or Project-centric) Enhance the compliance of capacitated individuals with supportive partners' (donors) accountability requirements	(Organizational Learning-centric)	(Developmental and Empowerment goal) Enhance change in countries' capacity to develop evaluation policies and strategies, commission evaluations, conduct data collection, disseminate and use evaluation findings by enhancing indigenous evidence-informed decision making and reducing countries' dependence and/or reliance on external technical support
<b>Main Premise</b>	(Assumptions are critical) Especially in the past those who implemented activities at this level believe individual in countries not to have sufficient evaluation capacity to attend to what are identified as critical roles and responsibilities. It was		(Context Analysis is essential) Countries already dispose Evaluation Capacity. It is the local actors' resources and ambitions that determine the capacity level but it is norms and policies that enhance the sustainability and ownership of evaluation

	external inputs and funding that created and fostered capacity. By focusing on individuals' training and awareness raising, organization- and government-wide change will take place		
<b>Key Outcomes</b>	<p>(Compliance Outcome)</p> <p>In the past, capacitated were considered individuals as “order takers.” Today, they are expected to promote even better evaluation designs and methodologies that their clients may call for</p> <p>Ensure the that the envisaged objectives of the projects which the capacitated individuals work on are achieved</p>	<p>(Process Enhancement Outcome)</p>	<p>(Mainstreaming Outcome)</p> <p>Enhance the set-up of highly participatory and contextually relevant processes in evaluation planning, implementation and use at the national level</p> <p>Intervention are designed based on an inter-sectorial diagnostic needs assessment (e.g., CAP SCAN conducted in the DRC in July 2010 to measure the evaluation and RBM capacity of 10 different ministries)</p>
<b>Key Objectives</b>	<p>(Individual-focused Approach or Reductionist approach)</p> <p>To enhancement evaluation competencies (knowledge and practice) among individuals selected within specific ministry/departments units or other segments of the organized sectors (project implementing units or community based organizations receiving donors' funding)</p>	<p>(Ensure improvement in attitude towards the evaluation)</p> <p>Ensure the availability of incentives and knowledge management processes towards a more consistent practice and use of evaluation within the organization</p>	<p>(Individual and Group-focused approach or Systemic approach)</p> <p>To enhance collective capabilities Integration of Evaluation <i>within country systems</i> (across all sectors – public, private and civil society - within society) often driven by a national evaluation policy and a national capacity development</p>

	<p>Ensure increase in knowledge about evaluation approaches, methods, managements and use</p> <p>Ensure improvements in evaluation related practices</p>		
<b>Focus</b>	<p>Doing and delivering (divergence from roadmap is penalized)</p> <p>Focus is to Supply a product</p>	<p>Reflection and learning (divergence from envisaged results is allowed and understood)</p> <p>Focus is to enable and increase use of evaluation</p>	<p>Maximize the results of the collective learning occurred at all three levels</p> <p>Focus is to Support the conduct and use of evaluation at all levels</p>
<b>Function</b>	<p>(Remedial function)</p> <p>To strengthen individuals' data collection and reporting skills</p> <p>Tactic</p>	<p>(Enabling function)</p> <p>To establish internal processes that might enable individuals' efforts in evaluation</p> <p>Incipient Strategy</p>	<p>(Propositional function)</p> <p>To pomote the sustainability of M&amp;E systems and institutional arrangements toward policy reforms</p> <p>Articulated Strategy</p>
<b>Scope and Modality of Delivery</b>	<p>Narrow</p> <p>Training/Workshop and technical assistance</p> <p>Project-based or Logframe-based approach)</p> <p>Classic Tailoring and technocratic customization</p> <p>Past Interventions at this level have often been designed based on the funders' program objectives rather than on the target project staff's capacity needs</p>	<p>Mixed</p>	<p>Broad</p> <p>Intermittent workshops combined with shadowing, coaching, mentoring and peer exchange (e.g. learning symposia or Communities of Practice)</p>

<p><b>Key Target Populations to capacitate</b></p>	<p>(Individuals located at the Macro and Micro Level)</p> <p><i>Macro Level:</i> Individuals within key ministries and governmental agencies</p> <p><i>Micro Level:</i> Key staff or contracted consultants working within NGOs and other non-institutional entities implementing funded projects</p> <p>(Atomized or “drop in the ocean” approach)</p> <p>Interventions are generally quite brief in duration, are rarely characterized by follow-ups and often involve a limited number of individuals scattered among agencies not related to each other.</p>	<p>Decision-makers or Opinion leaders within a variety of organizations</p> <p>Communities of Practice, networks of Civil society Organizations, Academia and Private Sector,</p>	<p>High-level policy-makers</p> <p>President’s Office and Parliament</p> <p>(Gradual, Incremental and opportunistic Process) Ministries and other governmental agencies</p> <p>Normally initiated in key government units (President or Prime Minister’s office or Ministry of Finance/Planning) as well as other specific sectors (Agriculture, Education and Health) and extended to other sectors at a later stage</p>
<p><b>Degree of Innovativeness</b></p>	<p>Innovativeness is not always the focus. Interventions at this level are not really questioning the status quo</p> <p>ECB welcomes the ideas that organizations and individuals may develop capacity over time but the focus is on facilitating functional improvements in the short- and medium-term</p>	<p>Innovativeness is sought after so long as it could enable the efforts of individuals within organizations to succeed in their evaluation endeavors.</p>	<p>Through a concerted effort, the focus is on improving countries’ welfare and promoting social change</p>

<b>Implementation Philosophy</b>	<p>Planning is based on predictions that, given activity <math>x</math> you will be able to achieve an outcome <math>x</math></p> <p>Aristotle Technical thought (following routing and a pre-set plan)</p> <p>Blueprint Approach</p>	<p>For individual capacity to develop, it is necessary to influence leadership and get the heads of the organizations on board. It is also important to create incentives from within and establish processes facilitating change</p>	<p>Planning consists in forecasting several scenarios and keeping flexible to address emerging needs</p> <p>Aristotle Practical thought – processes are disconfirmable, not self-seeking</p>
<b>Type of Accountability</b>	External	External and marginally internal	First of all, Internal. External, too, but not a priority
<b>Critics' argument</b>	<p>“Development as displacement”</p> <p>Donor centric financing and reporting displaced the purpose of producing global public goods that donors meant to support ECB is characterized by multiple idiosyncratic projects disconnected from overarching strategic objectives and metrics that disempower leadership (as well as by punctual capacity building activities organized around some general ministry strategic objectives</p>	<p>“Turn-over among organizational leaders is frequent”</p> <p>Organizations' leaders could drop out any time depending on the level of resources or time available, thus compromising the effectiveness and sustainability of evaluation</p>	<p>“ECD is too idealistic and/or ambitious”</p> <p>“ECD is too expensive”</p> <p>“ECD requires donors' harmonization but this is easier said than done”</p>
<b>Development</b>	Individual Empowerment Often Paternalistic Support In-	RBM and Aid Effectiveness Logic	MfDR and Development Effectiveness

<b>Perspective</b>	country entities as vulnerable Human Resource (skill development and training at the individual level)		Logic Resilience Booster: In-country entities are viewed as resilient and agents of change
<b>Equity and Vulnerability</b>	Unilateral transfer of knowledge, resources and technology	Adaptation and knowledge creation	ECD Exchange enabling mutual learning/benefit
<b>Link to the Public Sector Management Reform</b>	(Weak) Dedicated generous funding for M&E officers through externally funded projects' budgets Salary Top-ups	(Medium) Higher level of awareness of the repercussions on one's own work.	(Strong) Funding made available through governance (and not only evaluation budgets) Salaries of M&E officers in line with the national standards
<b>Paradigm of practice</b>	(Reductionist and confirmatory – it tests predictive capacity) Linear, oversimplified understanding of reality Emphasis on the “right answers” (OECD 2006)	Mixed	(Complex responsive and Adaptive) Emerging and Adaptive Goal-driven Emphasis on approaches that fit the context
<b>Most frequent design used to evaluate</b>	Pre- and Post-test  Mechanistic counting  Kirkpatrick Level 1 and 2	Kirkpatrick Level 3	Case Study  Tools: Self-assessment (IDRC, 2003) Semi-structured interviews Written Questionnaire  Kirkpatrick Level 4

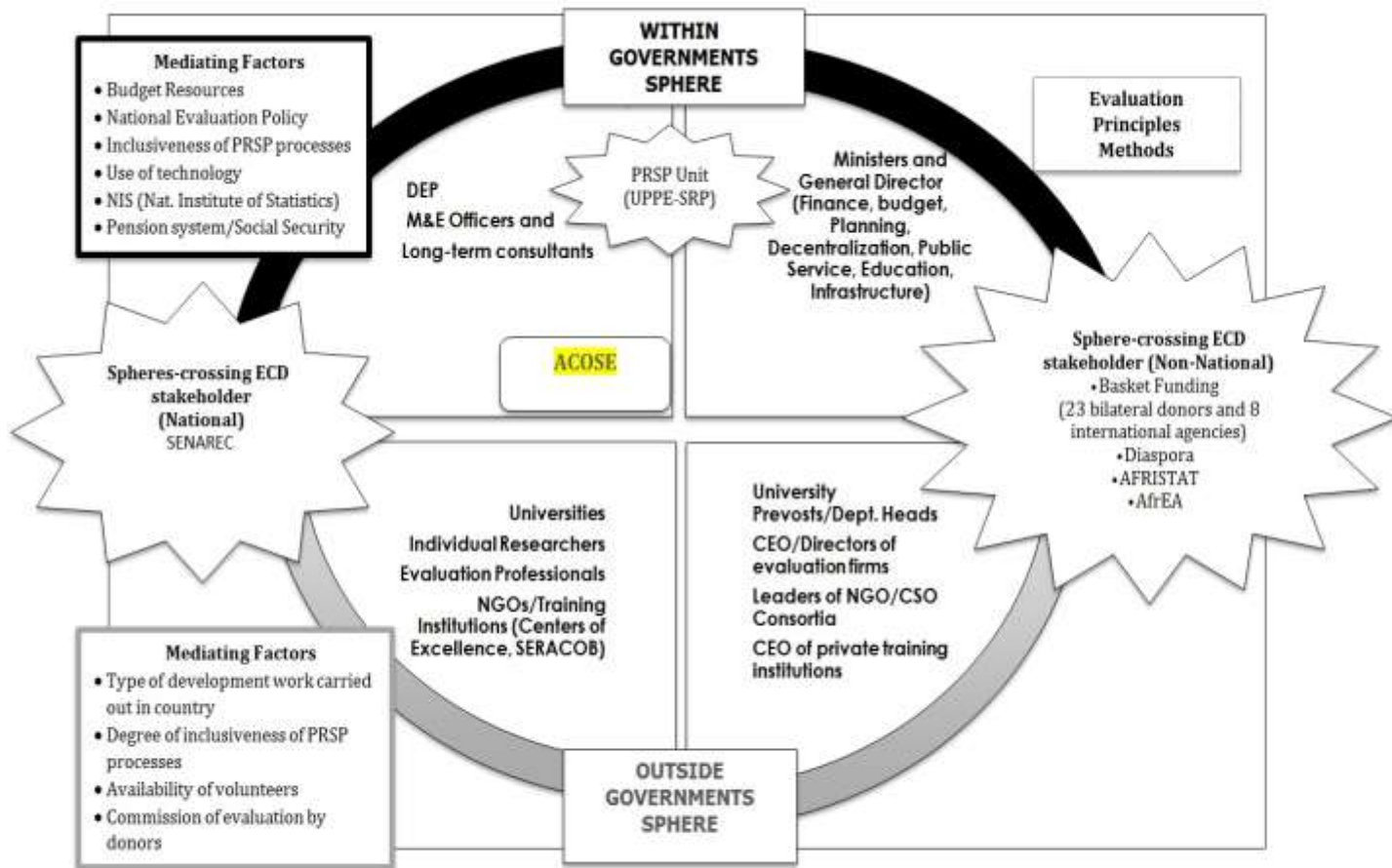


<b>Evaluation Champions</b>	A limited number of individuals turning over on a regular basis	Opinion leaders and decision-makers within organizations	Vertical Slice Championing: 1 or 2 Entities (rather than individuals) within the government  Diffused-championing  Long-term championing
<b>Personalization and location of the evaluation function</b>	High and concentrated (Evaluation is regarded as an activity)	High and More defused	Medium and more defused (Evaluation is institutionalized and regarded as a strategy)
<b>Availability of incentives</b>	Very limited incentives available	Increase in availability of incentives	During a preliminary phase, the value and contribution of each unit within the system is recognized through grants, tax exemptions and/or other types of facilitative measures (both monetary and non-monetary) towards both the production of better evaluation and the sustained use of evaluation findings. (OECD 2006)  However, over time, incentives are gradually suppressed
<b>Degree of systematic VOPE involvement in the design and implementation of ECD</b>	Low-Medium	Medium-High	High

<b>Funding source</b>	(External funding) Resources allocated by donors as stand-alone program or as an integral component of projects	Internal and External funding	(Mixed funding) Donors and national governments
<b>Nature of the Implementation Process</b>	A-political	Political	Relational and Political
<b>Ability to address Equity and Vulnerability</b>	Low	Contingent on the awareness level among opinion leaders and decision-makers	Medium-High
<b>Learning Objective</b>	Increase and improve competencies (Specific abilities of individuals)	Increase and improve capabilities (Specific abilities of organizational sub-systems)	Increase the creation and adaptation of normative rules acknowledging the relevance of the acquired capacities and capabilities. (Specific abilities of organizational sub-systems)

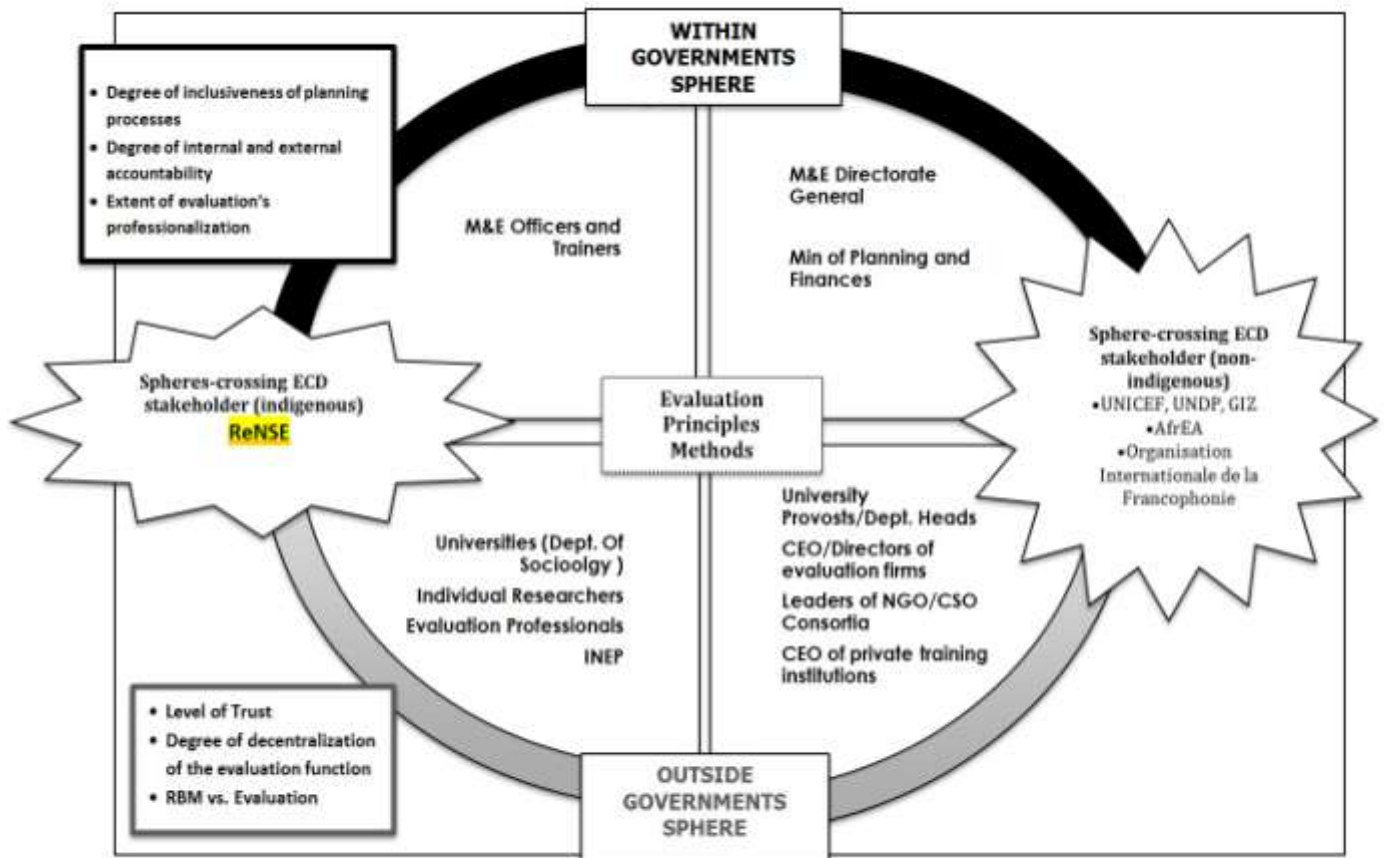
## Appendix K

### The SFAR Framework and the ECD Ecology in the DRC



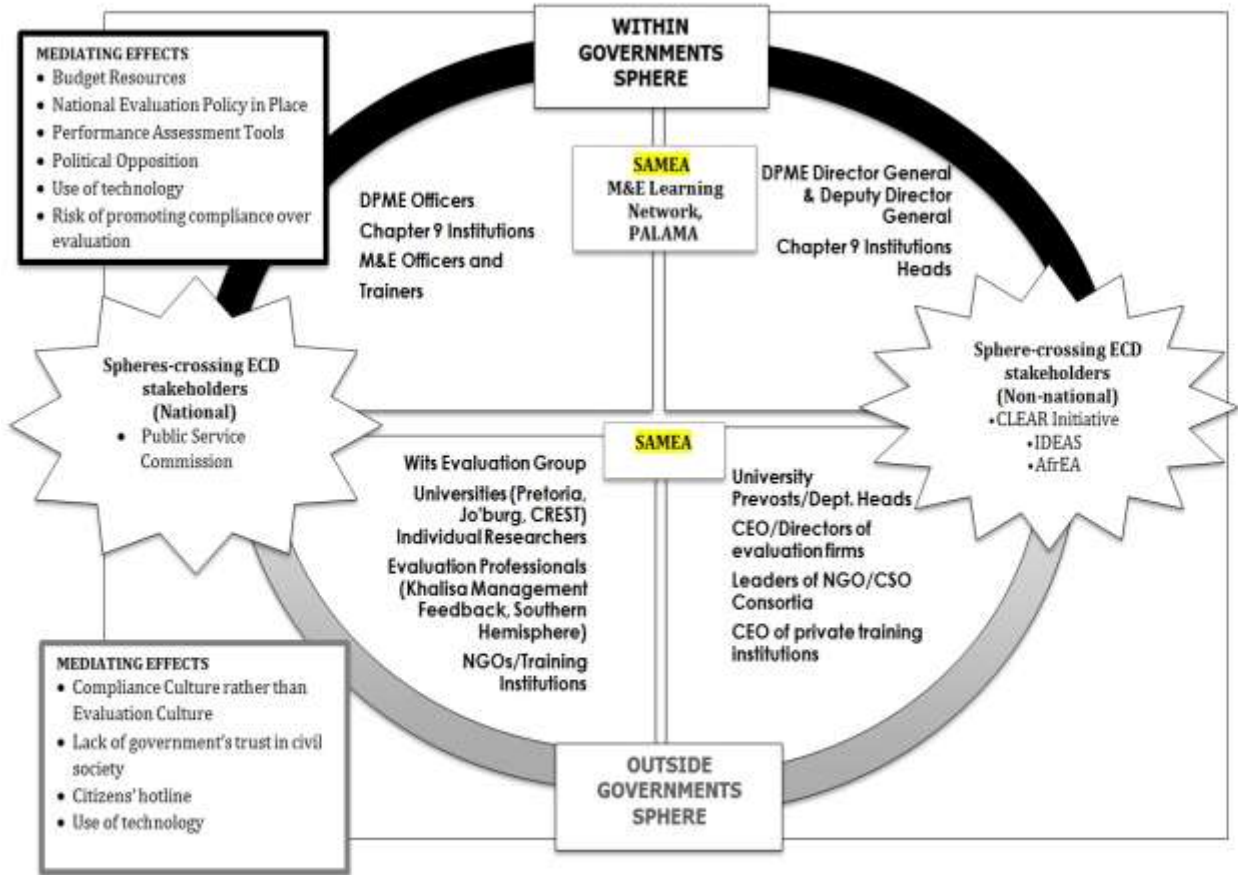
Appendix L

The SFAR Framework and the ECD Ecology in Niger



## Appendix M

### The SFAR Framework and the ECD Ecology in South Africa





## REFERENCES

- ACoSE. Rapport d'évaluation mi-parcours du Project BEED financé par USAID (composante mise en oeuvre par UNICEF) [Mid-term review of the USAID-funded BEED project component implemented by UNICEF]. (2011).
- African Development Bank & World Bank Operations Evaluation Departments (1998). *Evaluation Capacity Development in Africa: Selected Proceedings from a Seminar in Abidjan*. Tunis: African Development Bank.
- Altschuld, J.W., & Jumar D.D.(2010a). *Needs assessment: An overview*. Thousand Oaks, CA: Sage.
- Altschuld, J.W., & Eastmond, J. N. (2010b). *Needs assessment: Getting started*. Thousand Oaks, CA: Sage.
- Altschuld, J.W. (2010c). *Needs assessment: Data collection*. Thousand Oaks, CA: Sage.
- Altschuld, J.W., & White J. (2010d). *Needs assessment: Analysis and prioritization*. Thousand Oaks, CA: Sage.
- Altschuld, J.W., & King, J. (2010e). *Needs assessment: Taking action for global change*. Thousand Oaks, CA: Sage.
- Althsculd, J.W., & Witkin, B.R. (2000). *From needs assessment to action: Transforming needs into solution strategies*. Thousand Oaks, CA: Sage.
- Andrews, A.B., Motes, P.S., Floyd, A.G., Flerx, V.C., & Lopez-De Fede, A. (2006). Building evaluation capacity in community-based organizations: reflections of an empowerment evaluation team. *Journal of Community Practice*, 13, 85-104.
- Argyris, C. and Schön, D. (1996) *Organizational learning II: Theory, method and practice*. Reading, MA: Addison Wesley.
- Qureshi, S., Sweetser, A.T., & Buentjen, C. (2008). *Joint Study on Effective Technical Cooperation for Capacity Development: Pakistan Country Report*. Manila: Asian Development Bank
- Atkinson, D. D., Wilson, M., & Avula, D. (2005). A participatory approach to building capacity of treatment programs to engage in evaluation. *Evaluation and Program Planning*, 28, 329–334.
- Baizerman, M., Compton, D. W., & Stockdill, S. (2005). Capacity building. In S. Matthesion (Ed.), *Encyclopedia of evaluation* (pp. 38-39). Thousand Oaks, CA: Sage.
- Bakewell, O., & Garbutt, A. (2005). *The use and abuse of the logical framework approach*. Stockholm: Swedish International Development Cooperation Agency (SIDA).

- Bamberger, M. (Ed.). (2000). *Integrating quantitative and qualitative research in development projects*. Washington, DC: The World Bank
- Bamberger, M. (2006). Evaluation Capacity Building, In IOCE (2006). *Creating and developing evaluation organizations. Lessons learned from Africa, Americas, Asia, Australasia and Europe*. (pp.207-216). Lima: Desco.
- Bamberger, M. (2009). *Institutionalizing impact evaluation within the framework of a monitoring and evaluation system*. Washington, DC: The World Bank,
- Bamberger, M., Segone, M. (2011). *How to design and manage Equity-focused evaluations*. New York: UNICEF
- Bamberger, M., Rugh, J., & Mabry, L. (2012). *Real World Evaluation (2<sup>nd</sup> Ed.)*. Thousand Oaks, CA: Sage.
- Baser, H.; & Morgan, P. (2008). "Capacity, change and performance: Study report". In: ECDPM (Ed.), *Capacity, change and performance*. Maastricht: ECDPM.
- Baser H., Hauck V., & Land T. (2009). *Capacity development: Between planned interventions and emergent processes*. Maastricht: ECDPM.
- Bates, R.H., (1995). Social dilemmas and rational individuals: An assessment of the new institutionalism. In Harris, J., Hunter, J., & Lewis, C.L. (eds.), *Then new institutional Economics and third world development*. New York, NY: Routledge.
- Bawden, R. (2008). The educative purpose of higher education for human and social development in the context of globalization. In *Higher education in the world. Higher education: new challenges and emerging roles for human and social development*. Global University Network for Innovation. Basingstoke: Palgrave Macmillan.
- Bemelmans-Videc, M.L. (1992). Institutionalization evaluation: International Perspectives. In Mayne J. et al. (Eds.). *Advancing public policy evaluation: Learning from international experiences*. Amsterdam: Elsevier Science Publishers.
- Bemelmans-Videc, M.L., Rist, R.C., & Verdung E. (Eds). (1997). *Carrots, sticks and sermons: Policy instruments and their evaluation*. New Brunswick, NJ: Transaction Publishers.
- Berg, E. (Ed.). (1993). *Rethinking technical cooperation: Reforms for capacity building in Africa*, UNDP, New York.
- Bloom, P.N., & Dees, G. (2007). Cultivate your ecosystem. *Stanford Social Innovation Review*, 6, 47-53.
- Blue, R., Clapp-Wincek, C., & Benner, H. (2009). *Beyond success stories: Monitoring and evaluation for foreign assistance results*. Washington, DC: USAID

- Blue, R., Clapp-Wincek, C., & Benner, H. (2009). *Policy Brief: Monitoring and evaluation for foreign assistance results*. Washington, DC: USAID
- Boesen, N., & Therkildsen, O. (2005). A Results-Oriented Approach to Capacity Change. Copenhagen: Danida Evaluation Department (link: <http://www.um.dk/da/menu/Udviklingspolitik/MaalOgResultatstyring/Evaluering/OevrigeRapporter/CapacityDevelopment.htm>).
- Boesen, N. (2010). Looking for change beyond the boundaries, the formal and the functional. In J. Ubels, Na-A. Acquaye-Badoo, & A. Fowler, (Eds.), *Capacity Development in Practice* (pp.145-155) Earthscan: London.
- Bourgeois, I., Jill, A. C., & Cousins, J.B. (2008). Understanding organization capacity for evaluation: Synthesis and integration. *The Canadian Journal of Program Evaluation*, 23, 228-229.
- Boyle, R. (1995). *Towards a New Public Service*. Dublin: Institute of Public Administration.
- Boyle, R., Lemaire, D., & Rist, R. C. (1999). Incentives, constraints and culture-building as instruments for the development of evaluation demand. In R. Boyle & D. Lemaire (Eds.), *Building effective evaluation capacity* (pp. 153-177). New Brunswick, NJ: Transaction Publishing.
- Brandon, P. R., & Higa, A. F. (2004). An empirical study of building the evaluation capacity of K-12 site-managed project personnel. *Canadian Journal of Program Evaluation*, 19, 125-142.
- Britton, B. (2010). Monitoring and evaluation for personal learning. In J. Ubels, Na-A. Acquaye-Badoo, & A. Fowler, (Eds.), *Capacity Development in Practice* (pp. 264-276) Earthscan: London.
- Bruner, J. (1996). *The Culture of Education*, Cambridge, MA: Harvard University Press.
- Burton, P. (2009). *Conceptual, theoretical and practical issues in measuring the benefits of public participation*. Los Angeles, CA: SAGE.
- CDRA (2007) *Capacity Building: Critique and manifesto*. Cape Town: CDRA. Retrieved from <http://www.biennial.cdra.org.za/biennial-online/capacity-building-manifesto.html>
- Chambers, R. (1997). *Whose reality counts? Putting the first last*. London: Intermediate Technology.
- Checkland, P.B. (1999). *Soft Systems Methodology in Action*. Chichester: Wiley. Retrieved from <http://www.lums.lancs.ac.uk/profiles/peter-checkland/>
- Chelimsky, E. (2008). *A clash of cultures: Improving the "fit" between evaluative independence and the political requirement of a democratic society*. *American Journal of Evaluation*, 29, 400-415.

- Chung, K. (2000). Issues and approaches in the use of integrated methods. In M. Bamberger (Ed.), *Integrating quantitative and qualitative research in development projects* (pp.37-46). Washington DC: The World Bank.
- CIDA (2000). Capacity development. Why, what, how. *Capacity Development Occasional Series Vol. 1*, No. 1, May 2000. Ottawa: CIDA.
- Clotteau, G., Boily, M.-H., Darboe, S., & Martin, F. (2011). Building Capacities for Results-Based National M&E Systems. In Rist, R., et al. (Eds.) *Influencing Change: Building Evaluation Capacity to Strengthen Governance* (pp. 171-194). Washington, DC: The World Bank.
- Collier, P., & Chauvet, L. (2007). *What are the preconditions for turnarounds in failing states?* Oxford University and IRD, DIAL Retrieved from <http://users.ox.ac.uk/~econpco/research/pdfs/WhatPreconditionsForTurnaround.pdf>
- Compton, D., Baizerman, M., Preskill, H., Rieker, P., & Miner, K. (2001). Developing evaluation capacity while improving evaluation training in public health: The American Cancer Society's collaborative evaluation fellows project. *Evaluation and Program Planning*, 24, 33-40.
- Compton, D. W., Baizerman, M., & Stockdill, S. H. (Eds.). (2002). The art, craft, and science of evaluation capacity building: *New directions for evaluation*, 93, 1-120. San Francisco: Jossey-Bass.
- Compton, D. W., Glover-Kudon, R., Smith, I. E., & Avery, M. E. (2002). Ongoing capacity building in the American Cancer Society (ACS) 1995–2001. In D. W. Compton, M. Baizerman, & S. H. Stockdill (Eds.), *The art, craft and science of evaluation capacity building*. *New Directions for Evaluation*, 93, 47–61.
- Constantinou, N. (2007). Capacity Development in the World Bank Group: A review of non-lending approaches. *Capacity Development Briefs*. No.23. Washington, DC: The World Bank.
- Coryn, C.L.S. (2007). *Evaluation of researchers and their research: Toward making the implicit explicit*. Unpublished doctoral dissertation, Western Michigan University, Kalamazoo.
- Cousins, J.B., & Whitmore, E. (1998). Framing Participatory Evaluation. *New Directions for Evaluation*, 80, 5-23.
- Cousins, J.B. (2003). Utilization efforts of participatory evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 245-266). Dordrecht, Great Britain: Kluwer Academic Publishers.
- Cousins, J.B., Goh, S.C., Clark, S., & Lee, L.E. (2004). Integrating evaluative Inquiry into the organizational culture: A review synthesis of the knowledge base. *Canadian Journal of Program Evaluation*, 19, 99-141.

- Cracknell, B.E. (2000). *Evaluating development aid: issues, problems and solutions*. Thousand Oaks, CA: Sage.
- Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.
- Crutchfield, L.R., & Grant, H.M. (2008). *Forces for good: the six practices of high-impact nonprofits*. San Francisco, CA: Jossey-Bass.
- Dabelstein, N. (2003). Evaluation capacity development: Lessons learned. *Evaluation*, 9, 365–369.
- Dahlgren, S. (2011). Lessons Learned in Capacity Building: Where Do We Go from Here?. In Rist et al. (Eds.), *Influencing change: Building evaluation capacity to strengthen Governance* (pp.243-261). Washington, DC: The World Bank
- Dart, J. J. & Davies, R.J. (2003). A dialogical story-based evaluation tool: The most significant change technique, *American Journal of Evaluation* 24, 137–155
- Datta, L. (1997). Multimethod evaluations: Using case studies together with other methods. In Chelimsky, E. & Shadish, W. (Eds.), *Evaluation for the 21<sup>st</sup> century*. Thousand Oaks, CA: Sage.
- DeGEval (2012). *Standards für evaluation*. [German Evaluation Society Standards]. Retrieved from HYPERLINK "<http://www.degeval.de/degeval-standards/standards>" <http://www.degeval.de/degeval-standards/standards>
- Derlien, H.U. (1990). Genesis and structure of evaluation efforts in comparative perspective. In Rist, R. (Ed.), *Program evaluation and the management of government: Patterns & prospects across eight nations*. New Brunswick, NJ: Transaction Publishers.
- Dillman, D. (2000). *Mail and Internet Surveys: A tailored design method*. New York: Wiley.
- Douglah, M., Boyd, H., & Gundermann, D. (2003, November). *Nurturing the development of an evaluation culture in a public educational agency*. Paper presented at the annual meeting of the American Evaluation Association, Atlanta, GA.
- DPME (2012, April). Presentation delivered by DPME staff at a conference on country M&E systems in Africa, sponsored by the Clear Initiative in Pretoria.
- Duffy, J.L., & Wandersman, A. (2007, November). *A review of research on evaluation capacity-building strategies*. Paper presented at the annual conference of the American Evaluation Association, Baltimore, MD.
- Durlak, J.A., & DuPre, E.P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327-350.

- Earl, S., Carden,, F., & Smutylo, T. (2001). *Outcome mapping. Building learning and reflection into development programs*. Ottawa: International Development Research Center.
- Easterley, W. (2007). *The White man's burden: Why the West's efforts to aid the rest have done so Much ill and so little good*. London: Penguin Press.
- Ellerman, D. (2001). Helping people help themselves: Toward a theory of Autonomy-Compatible Help”. *Policy Research Working Paper No 2693*. The World Bank: Washington DC.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *The Handbook of research on teaching*. New York: MacMillan.
- EuropeAid (2008). *Analysing and addressing governance in sector operations*. Brussels: EuropeAid.
- ECDG (2012). *The New ECDG Guide to Evaluation Capacity Development based on the IWA on ECD*. Geneva: ECDG.
- Engel, P., Keijzer, N., Land, T. 2007. A balanced approach to monitoring and evaluating capacity and performance: A proposal for a framework. *ECDPM Discussion Paper No. 58E*. Maastricht: ECDPM. Retrieved from [http://www.ecdpm.org/Web\\_ECDPM/Web/Content/Content.nsf/vwPrint/A1649F5CBFFEB410C12573B5005C5030?Opendocument](http://www.ecdpm.org/Web_ECDPM/Web/Content/Content.nsf/vwPrint/A1649F5CBFFEB410C12573B5005C5030?Opendocument)
- Eyben, R., Kidder, T., Rowlands, J., & Bronstein, A. (2008). Thinking about change for development practice: A case study for Oxfam GB. *Development in Practice, 18*, 12.
- Fetterman, D., & Wandersman, A. (Eds.). (2005). *Empowerment evaluation principles in practice*. New York, NY: Guildford Press.
- Firestone, W. A. (1993). Alternative arguments for generalizing from data as applied to qualitative research. *Educational Researcher, 22*, 16-23.
- Fleischer, D., Christie, C. A., & LaVelle, K. B. (2011). Perceptions of evaluation capacity building in the U.S.: A descriptive study of American Evaluation Association members. *Canadian Journal of Program Evaluation, 23*, 37–60
- Flyvbjerg, B. (2006). Five misunderstanding about case-study research. *Qualitative Inquiry 12*, 2, 219-245
- Fowler, A. (2008). Connecting the dots: Complexity thinking and social development. *In The Broker Online*. Retrieved from <http://www.thebrokeronline.eu/en/Articles/Connecting-the-dots>
- Fowler, A., & Ubels, J. (2010).The multifaceted nature of capacity: Two leading models. In J. Ubels, Na-A. Acquaye-Badoo, & A. Fowler, (Eds.), *Capacity Development in Practice* (pp.11-24) Earthscan: London.

- Freire, P. (1970). *Pedagogy of the Oppressed*. London: Continuum Publishing.
- Fukuda-Parr, S., Lopes, C., & Malik K. (Eds.). (2002). *Capacity for development: New solutions to old problems*. New York: UNDP.
- Fukuyama, F. (2004). *State-building: Governance and world order in the 21st century*. Ithaca, NY: Cornell University Press.
- Fung, A. (2003). Recipes for public spheres: Eight institutional design choices and their consequence. *Journal of Political Philosophy*, 11, 338–67.
- Furubo, Rist, & Sandahl (Eds.) (2002). *International atlas of evaluation* (pp.7-9). New Brunswick, NJ: Transaction Publishers.
- Gaventa, J. (2005). Reflections of the uses of the ‘power cube’ approach for analyzing the spaces, places and dynamics of civil society participation and engagement. *CFP Evaluation Series No 4*. The Hague: MBN. Retrieved [HYPERLINK "http://www.hivos.nl/index.php/content/download/6134/34591/file/da744d27f6309b31f60e1c508a%201b4a0b.pdf"](http://www.hivos.nl/index.php/content/download/6134/34591/file/da744d27f6309b31f60e1c508a%201b4a0b.pdf)  
[www.hivos.nl/index.php/content/download/6134/34591/file/da744d27f6309b31f60e1c508a1b4a0b.pdf](http://www.hivos.nl/index.php/content/download/6134/34591/file/da744d27f6309b31f60e1c508a1b4a0b.pdf)
- Gaventa, J., & Barrett, G.(2010). So what difference does it make? Mapping the outcomes of citizen engagement. *IDS Working Paper 347*. Brighton, UK: Institute of Development Studies.
- German Development Cooperation (2012). *Monitoring and Evaluation at GIZ*. Retrieved from [http://www.giz.de/en/aboutgiz/monitoring\\_and\\_evaluation.html](http://www.giz.de/en/aboutgiz/monitoring_and_evaluation.html)
- Gervais, M., & Johnson, H. (2007). Une experience de collaboration en Afrique [A story of collaboration in Africa]. *Bulletin de la Société Québécoise de l'Évaluation*, 19, 14-17.
- Gibbs, D., Napp, D., Jolly, D., Westove, B., & Uhle, G. (2002). Increasing evaluation capacity within community-based HIV prevention programs. *Evaluation and Program Planning* 25, 261-269
- Gilliam, A., Barrington, T., Davis, D., Lacson, R., Uhl, G., & Phoenix, U. (2003). Building evaluation capacity for HIV prevention programs. *Evaluation and Program Planning*, 26, 133-142.
- Gray, A., Jenkins, B., & Segsworth, B., (Eds.). (1993). *Budgeting auditing and evaluation: Functions and integration in seven governments*. New Brunswick, NJ: Transaction Publishers.
- Greene, J.C. (2005). Stakeholder involvement. In S. Mathison (ed.), *Encyclopedia of Evaluation* (p. 397). Thousand Oaks, CA: Sage.

- Greene, J.C., & Caracelli, V.J., (Eds.) (1997). Advances in mixed method evaluation: The challenges and benefits of integrating diverse paradigms. *New Directions for Evaluation*. No.74. San Francisco: Jossey-Bass.
- Guba E., & Lincoln, Y. (1989) *Fourth Generation Evaluation*. Newbury Park, CA: Sage.
- Guijt, I. (2008). *Assessing and learning for social change. Discussion paper*. Brighton: Institute of Development Studies.
- Guerrero O. R. (1999). Comparative insights from Columbia, China and Indonesia. *Evaluation Capacity Development Working Paper Series No. 5*. Washington DC: The World Bank.
- Havens, H.S. (1992). The Erosion of Federal Program Evaluation. In C.G. Wye, & R.C. Sonnichsen (Eds.). *New Directions for Program Evaluation*, 55.
- Hay, K. (2010). Evaluation field building in south Asia: Reflections, anecdotes, and questions. *American Journal of Evaluation*. 31, 222.
- Heider, C. (2011). A Conceptual Framework for Developing Evaluation Capacities: Building on Good Practice. In Rist et al. (Eds.), *Influencing Change: Building Evaluation Capacity to Strengthen Governance* (pp. 95-100). Washington DC: The World Bank.
- Henry, G.T., & Mark, M.M. (2003). Beyond Use: Understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation* 24, 293-314.
- Hettlage, R., & Steinlin, M. (2006). *The critical incident technique in knowledge management-related contexts*. Zurich, Switzerland: Ingenious-Peoples-Knowledge. Retrieved from [http://www.click4it.org/images/6/68/Critical\\_IncidentTechnique\\_in\\_KM.pdf](http://www.click4it.org/images/6/68/Critical_IncidentTechnique_in_KM.pdf)
- Hildebrand, Mary E. and Merilee S. Grindle (1994). *Building capacity: Challenges for the public sector*. Cambridge, MA: Harvard University.
- Higa, T. A. F., & Brandon, P. R. (2008). Participatory evaluation as seen in a Vygotskian framework. *Canadian Journal of Program Evaluation*, 23, 103–125
- Himmelman, A. (1996). On the theory and practice of transformational collaboration: from social service to social justice. In C. Huxham (Ed.) *Creating Collaborative Advantage* (pp.19-43). London: Sage.
- Himmelman, A. T. (2001) On Coalitions and the transformation of power relations: Collaborative betterment and collaborative empowerment. *American Journal of Community Psychology*, 29, 227-235.
- Hoeffler, A., & Outram, V. (2008) *Need, merit or self-interest: What determines the allocation of aid?* Oxford: CSAE



- Holton, E.F., Bates, R.A., & Naquin, S.S. (2000). Large-scale performance-driven training needs assessment: A case study. *Public Personnel Management, 29*, 249-265.
- Holvoet, N., Dewachter, S., & Gildemyn, M. (2011). Strengthening national evaluation societies: An exploratory survey. Maastricht: Institute for Development Policy and Management.
- Hoole, E. & Patterson, T.E., (2008). Voices from the field: Evaluation as part of a learning culture. *New Directions for Evaluation, 119*, 93-113.
- 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).
- House, E., & Howe., K.R. (1999). *Values in Evaluation and Social Research*. Thousand Oaks, CA: Sage.
- Huffman, D., Thomas, K., & Lawrenz, F. (2008). A Collaborative Immersion Approach to Evaluation Capacity Building. *American Journal of Evaluation, 29*, 358-368.
- IDEAS (2011). *Guidelines on key competencies for development evaluators, managers and commissioners*.
- IFPRI (2011). Global Hunger Index. Washington, DC. IFPRI.
- IOCE (2006). *Creating and developing evaluation organizations. Lessons learned from Africa, Americas, Asia, Australasia and Europe*. Lima: Desco.
- Jalijouli, M.A. (2011) Evaluation systems as strategy management tools: building Dubai's institutional learning capacity. In Rist et al. (Eds.) *Influencing Change: Building Evaluation Capacity to Strengthen Governance* (pp.59-84). Washington: DC The World Bank.
- James, R. (2001) Practical Guidelines from the Monitoring and Evaluation of Capacity Building: Experiences from Africa. In INTRAC (Ed.) *Occasional Paper Series*. Oxford, INTRAC.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluation of educational programs* (2<sup>nd</sup> ed.). Thousand Oak, CA: Sage.
- Khan, M. A. (1998). Evaluation capacity building: An overview of current status, issues and options. *Evaluation, 4*, 310–328.
- Kaplan, A. 1999. *The developing of capacity*. Cape Town: Community Development Research Center.

- Kaufman, R., Oakley-Brown, H., Watkins, R., & Leigh, D. (2003). *Strategic planning for success: Aligning people, performance, and payoffs*. San Francisco: Jossey-Bass.
- King, J. A. (2002). Building the evaluation capacity in a school district. In D. W. Compton, M. Baizerman, & S. Stockdill (Eds.). *The art, craft, and science of evaluation capacity building*. *New directions for evaluation*, 93, 63–80.
- King, J. A., & Volkov, B. (2005). A framework for building evaluation capacity based on the experiences of three organizations. *CURA Reporter*, 35, 10–16.
- Kirkpatrick, D. (1998). *Evaluating Training Programs: The Four Levels*. San Francisco: Berrett-Koehler.
- Krippendorff, K. (1986). *Information theory: Structural models for qualitative data*. Thousand Oaks, CA: Sage.
- Kurtz, C., & Snowden, D., (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM Journal of Research and Development*, 42, 462-483.
- Kusek, J., & Rist, R. (2004). *Ten steps to a results-based monitoring and evaluation system: A handbook for development practitioners*. World Bank: Washington DC.
- Kuzmin, A. (2004). *Exploration of factors that affect the use of evaluation training in evaluation capacity development*. Unpublished doctoral dissertation, Union Institute & University, Cincinnati, Ohio.
- Labin, S.N., Duffy, J.F., Meyers, D.C., Wandersman, A., & Lesene, C.A. (2012). A Research Synthesis of the Evaluation Capacity Building Literature. *American Journal of Evaluation*, 33, 307-338.
- 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.
- Lahiri, S. (2007). *Theory and Practice of Foreign Aid*. London, UK: Elsevier Science.
- Land, T., Hauck, V., & Baser, H. (2009). Capacity development between planned interventions and emergent processes: Implications for development cooperation. *Policy Management Brief*. No. 22.
- Lavergne, R., Lewis, E. & Pecore-Ugorji, D. (2004). *Capacity Development in CIDA's Bilateral Programming: A Stocktaking*. Gatineau: CIDA Policy Branch (draft).
- Leach, M., & Scoones, I. (2007). Mobilizing citizens: Social movements and the politics of knowledge. *IDS Working Paper 276*. Brighton: IDS.
- Lee, Y.S. (1999). Evaluation Coverage in R. Boyle & D. Lemaire (Eds.), *Building effective evaluation capacity* (pp. 153-177). New Brunswick, NJ: Transaction Publishing.

- Leeuw, F.L., Rist, R.C., & Sonnichsen R.C. (Eds.) (1994), *Can government learn? Comparative perspectives on evaluation and organizational learning*. New Brunswick, NJ: Transaction Publishers.
- Lennie, J. (2005). An Evaluation Capacity-Building Process for Sustainable Community IT Initiatives: Empowering and Disempowering Impacts. *Evaluation, 11*, 390-414.
- Lessik, A., & Michener, V. (2000). Measuring institutional capacity. In USAID CDIE (Ed.), *Recent Practices in Monitoring and Evaluation TIPS*. Washington, DC: USAID.
- Leviton, L. C. (2001). Presidential address: Building evaluation's collective capacity. *American Journal of Evaluation, 22*, 1-12.
- Lewis, C., Perry, R., & Murata, A. (2006). How Should Research Contribute to Instructional Improvement? The Case of Lesson Study. *American Journal of Evaluation, 35*, 3-14.
- Lincoln, Y.S., & Guba, E.G . (1985). *Naturalistic Enquiry*. Beverly Hills, CA: Sage.
- Liverani, A., & Lundgren, H. E. (2007). Evaluation systems in development aid agencies. An analysis of DAC Peer Reviews 1996 -2004. *Evaluation 13*, 241-256.
- Love, A., (2006). Developing evaluation capacity in extension 4-H field faculty: A framework for success. *American Journal of Evaluation, 22*, 257-269.
- Lusthaus, C, Adrien, M.H., Anderson, G., Carden, F. & Montalvan G.P. (2002). *Organizational assessment. A framework for improving performance*. Ottawa/Washington DC: International Development Research Centre and Inter-American Development Bank.
- MacKay, K. (1999). A diagnostic guide and action framework. In *ECD Working Paper Series, 6*, January 1999. Washington DC: World Bank. Retrieved from [http://siteresources.worldbank.org/EXTEVACAPDEV/Resources/4585672-1252437578683/ECD\\_WP6.pdf](http://siteresources.worldbank.org/EXTEVACAPDEV/Resources/4585672-1252437578683/ECD_WP6.pdf)
- Magruder-Watkins, J., & Mohr, B. (2001). *Appreciative inquiry: Change at the speed of imagination*. New York, NY: Jossey-Bass
- Maher, C.A. (1981). Training of managers in program planning and evaluation: Comparison of two approaches. *Journal of Organizational Behavior Management 3*, 45-56.
- Mayne, J. (2008) *Building an evaluative culture for effective evaluation and results management*. ILAC Working Paper 8. Rome: Institutional Learning and Change Initiative.
- Mayne, J., and Zapico-Goñi (1997). *Monitoring Performance in the Public Sector: Future Direction from International Experience*. New Brunswick, NJ: Transaction Publishers.

- McAllister, E. J., (2011). Old Challenges and New Frontiers. In Rist, R. et al. (Eds.), *Influencing Change: Building Evaluation Capacity to Strengthen Governance* (pp. 207-224). Washington DC: The World Bank.
- McAllister, E. J., Bezanson, K., Chandra, G., & Waage, J. (2008). *Bringing Together the Best of Science and the Best of Development*. Report to the Executive Council of the Consultative Group for International Agricultural Research (CGIAR) by the CGIAR Independent Review Panel. Retrieved from <http://www.cgiar.org/externalreview/>.
- McDonald, B., Rogers, P., & Kefford, B. (2003). Teaching people to fish? Building evaluation capacity of public sector organizations. *Evaluation, 9*, 9-29.
- McLeod, S. A. (2007). *Vygotsky Social Development Theory*. Retrieved from <http://www.simplypsychology.org/vygotsky.html>
- Mehta, L., Thompson, L., & Nleya, N. (2010). Claiming Water Rights in India and South Africa. In L. Thompson & C. Tapscott (Eds.), *Citizenship and social movements: Perspectives from the global south*. New York, NY: Zed
- Mertens, D.M. (2009). *Transformative Research and Evaluation*. New York, NY: The Guilford Press.
- Miles, M.B., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook*, Thousand Oaks, USA: Sage Publications.
- Milstein, B., Chapel, T. J., Wetterhall, S.F., & Cotton, D. A. (2002). Building capacity for program evaluation at the Centers for Disease Control and Prevention. *New Directions for Evaluation, 93*, 27-46.
- Missika–Wierzba, B., & Nelson, M. (2006). A revolution in capacity development? Africans ask tough questions. *Capacity Development Briefs No.16*. Washington DC: The World Bank.
- Moore, M. (1995). Promoting Good Government by Supporting Institutional Development. *IDS Bulletin, 26*, 2.
- Morazán, P., & Knoke, I. (2005) PRSP is about people: *Experiences and reflections with partner organisations in five African countries (Cameroon, DR Congo, Ethiopia, Ghana and Tanzania)*. Berlin: Brot für die Welt, Miseror, Evangelischer Entwicklungsdienst.
- Morell, J.A. (2010). *Evaluation in the face of uncertainty: Anticipating surprise and responding to the inevitable*. New York, NY: Guildford.
- Morgan, P. (1998). *Capacity and capacity Development: Some strategies*. Hull: Policy Branch, CIDA.
- Morgan, P. (2005a). *The Idea and Practice of Systems Thinking and their Relevance for Capacity Development*. European Centre for Development Policy Management: Maastricht, The Netherlands.

- Morgan, P. (2005b). Organising for large-scale system change: The environmental action (ENACT) programme, Jamaica. A case study for the project 'capacity, change and performance. *ECDPM Discussion Paper 57*. Maastricht: ECDPM. Retrieved <http://www.ecdpm.org/dp57>
- Morgan, P. (2006) *The concept of capacity: Capacity, change and performance*. Maastricht: European Centre for Development Policy Management.
- Morra Imas, L.G., & Rist, R.C. (2009). *The road to results: Designing and conducting effective development evaluations*. Washington, DC: World Bank.
- Morris, M. (1994). The Role of single evaluation courses in evaluation training. *New Directions of Program Evaluation*, 62, 17-27.
- Moseley, J.L., & Heaney, M.J. (1994). Needs assessment across the disciplines. *Performance Improvement Quarterly*, 7, 60-79.
- Mouton, C. (2010). *The history of program evaluation in South Africa*. Unpublished master's thesis. University of Stellenbosch.
- Mowles, C., Stacey, R., Griffin, D. (2008). What contribution can insights from the complexity sciences make to the theory and practice of development management? *Journal of International Development*, 20, 804-820.
- Newcomer, K. E. (2004). How might we strengthen evaluation capacity to manage evaluation contracts? *American Journal of Evaluation*, 25, 209-218.
- Nielsen, S. B.; & Attström, K. (2011). Perspectives on evaluation capacity building. In Rist, R. et al. (Eds.), *Influencing Change: Building Evaluation Capacity to Strengthen Governance*. (pp. 225-242). Washington, DC: The World Bank.
- Nielsen, B., Lemire, S., & Skov, M. (2011). Measuring Evaluation Capacity – Results and Implications of a Danish Study. *American Journal of Evaluation* 32, 324-344
- Ntalaja, K. (2005, December). Implications of civil society involvement in PRSP processes in DRC. Paper presented at the Civil Society Forum on “Sustainable Development in the DRC: Citizens’ participation and National and International opportunities”. Paper presented at the Nganda Catholic Center.
- Norwegian Agency for Development Cooperation (2006). *Instruction for evaluation activities in Norwegian aid administration*. Retrieved from <http://www.norad.no/en/evaluation>
- Ramalingam, B., Jone, H., Reba, T., & Young, J., ODI. (2008). *Exploring the science of complexity: Ideas and implications for development and humanitarian efforts*. ODI's working paper no. 235.
- OECD (2005). The Paris Declaration on Aid Effectiveness. Paris. Retrieved from <http://www.oecd.org/dataoecd/11/41/34428351.pdf>

- OECD (2006a). *The challenges of capacity development: Working towards good practice*. Paris.  
Retrieved from  
<http://www.oecd.org/dac/governanceanddevelopment/36326495.pdf>
- OECD (2006b). *Survey on monitoring the Paris Declaration: Overview of the results*. Paris.
- OECD (2008). Accra Agenda for Action. Retrieved from  
<http://www.oecd.org/dataoecd/11/41/34428351.pdf>
- OECD (2010). *Towards a strategic approach for evaluation capacity development*. Draft note prepared for the OCED 11th Meeting.
- OECD (2011a). DAC Criteria for Evaluating Development Assistance. Retrieved from  
HYPERLINK  
"[http://www.oecd.org/document/22/0,2340,en\\_2649\\_34435\\_2086550\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/22/0,2340,en_2649_34435_2086550_1_1_1_1,00.html)"  
[http://www.oecd.org/document/22/0,2340,en\\_2649\\_34435\\_2086550\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/22/0,2340,en_2649_34435_2086550_1_1_1_1,00.html)
- OECD (2011b). Development evaluation resources and systems: A study of network members. Retrieved from HYPERLINK  
"<http://www.oecd.org/development/evaluationofdevelopmentprogrammes/dcdndep/45605026.pdf>"  
<http://www.oecd.org/development/evaluationofdevelopmentprogrammes/dcdndep/45605026.pdf>
- OECD (2012). *Statistics on resource flows in developing countries*.  
<http://www.oecd.org/dac/aidstatistics/statisticsonresourceflowstodevelopingcountries.htm>
- Ortiz, A., & Taylor P. (2008). *Learning purposefully in capacity development: Why, when and what to measure?* Opinion Paper for IIEP. Brighton, Sussex: Institute of Development Studies (IDS),
- Otoo, S., Agapitova, N., & Behrens J. (2009). *The capacity development results framework: A strategic and results-Oriented approach to learning for capacity development*. Washington D.C.: World Bank Institute.
- O'Sullivan, R.G. (2004). *Practicing evaluation: A collaborative approach*. Thousand Oaks, CA: Sage.
- Owen, J. M. (2003). Evaluation culture: A definition and analysis of its development within organizations. *Evaluation Journal of Australasia*, 3, 43-47.
- Pasteur, K. 2006. Learning for development. In Eyben R., (Ed.), *Relationships for aid*. London: Earthscan.

- Patton, M.Q. (2003). *Qualitative Evaluation Checklist*. Kalamazoo, MI: Western Michigan University. Retrieved from [http://www.wmich.edu/evalctr/archive\\_checklists/qec.pdf](http://www.wmich.edu/evalctr/archive_checklists/qec.pdf)
- Patton, M.Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York, NY: Guildford.
- Piaget, J. (1995). *Sociological Studies*. London: Routledge.
- Picciotto, R., (1995). Introduction: Evaluation and development. *New Directions for Evaluation*, 3, 13-23.
- Picciotto, R., (1998). Evaluation capacity development: Issues and challenges. In *Evaluation Capacity Development in Africa: Selected Proceedings from a Seminar in Abidjan*, (pp.39-46). Washington DC: The World Bank.
- Picciotto, R. (2011a). *The Logic of Evaluation Professionalism*. *American Journal of Evaluation*, 17, 165.
- Picciotto, R. (2011b). Where is development evaluation going?. In Rist R. et al. (Eds.), *Influencing change: Building evaluation capacity to strengthen governance* (pp. 195-206). Washington, DC: The World Bank.
- Piper, L. & Nadvi, L. (2010). Popular mobilization, party dominance and participatory governance in South Africa. In L. Thompson and C. Tapscott (Eds.), *Citizenship and Social Movements: Perspectives from the Global South*. New York, NY: Zed.
- Podems, D. (2007, June). *Gender and Feminist Issues in Global Evaluation*. Paper presented at the AEA/CDC Summer Institute, Atlanta, Georgia.
- Porter, S. (2011). Recognizing “helping” as an evaluation capacity development strategy. In Rist R. et al. (Eds.), *Influencing change: Building evaluation capacity to strengthen governance* (pp.151-170). Washington, DC: The World Bank.
- Potter C., & Brough R. (2004). *Systemic capacity building: A hierarchy of needs*. Oxford, UK: Oxford University Press.
- Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, 29, 443-459.
- Preskill, H., & Boyle, S. (2008). Insights into evaluation capacity building: Motivations, strategies, outcomes, and lessons learned. *The Canadian Journal of Program Evaluation*, (3), 147.
- Preskill, H., & Torres, R., (2009). Readiness for organizational learning and evaluation (ROLE). In D. Russ-Eft & H. Preskill (Eds.), *Evaluation in organizations* (2<sup>nd</sup> ed.)(pp. 491-504). Boston, MA: Perseus Books.

- Rademacher, A. (2005). The growth of capacity in IUCN in Asia. In *ECDPM Discussion Paper 57M*. Maastricht: ECDPM Retrieved from HYPERLINK "http://www.ecdpm.org/dp57" http://www.ecdpm.org/dp57
- Ramage, M., & Shipp, K. (2009). *System thinkers*. London: Springer.
- Ramalingam, B., Jones, H., Reba, T., & Young, J. (2008). Exploring the science of complexity: Ideas and implications for development and humanitarian efforts. HYPERLINK "http://www.odi.org.uk/publications/working\_papers/WP285.pdf" *ODI Working Paper*, 285. London: ODI.
- Reeler, D., (2007). *A Theory of social change and implications for practice, planning, monitoring and evaluation*. Cape Town: Community Development Resource Association.
- Reynolds, M. & B. Williams (2012). *Systems thinking and equity-focused evaluations* (pp.. In Bamberger and Segone (2012). How to design and manage Equity-focused evaluations. New York: UNICEF
- Rist R. et al. (Eds.). (2011). *Influencing change: Building evaluation capacity to strengthen governance*. Washington, DC: The World Bank.
- Rist, R., & Stame, N. (Eds.)(2006). *From studies to streams: Managing evaluative systems*. New Brunswick, New Jersey: Transaction Book.
- Robinson, T.T., & Cousins, J.B. (2004). International participatory evaluation as organizational learning: A longitudinal case study. *Studies in Educational Evaluation*, 30, 1-22.
- Rocha Menocal, A., & Sharma, B. (2008). *Joint evaluation of citizens' voice and accountability: Synthesis Report*. London: Department for International Development.
- Rodríguez-Bilella, P., & Monterde-Díaz, R. (2010). Evaluation, valuation, negotiation: Some reflections towards a culture of evaluation. *The Canadian Journal of Program Evaluation*, 25, 1.
- Rodriguez-Campos, L. (2005). *Collaborative evaluations*. Tamarac, FL: Llumina Press.
- Rogerson, A., & Steensen, S. (2009) Aid Orphans: Whose responsibility? *OECD Development Co-operation Directorate Development Brief*. Paris: OECD
- Rugh, J. (2012, November) *The amazing growth of the global evaluation profession. Summary of findings from the EvalPartners mapping survey*. Paper presented at the American Evaluation Association Annual Meeting in Minneapolis.
- SAMEA (2010). *Membership Directory*.
- Sanders, J.R. (2003). Mainstreaming evaluation. *New Direction of Evaluation*, 9, 3-6.



- Schaumberg-Muller, H. (1996). *Evaluating capacity building : Donor support and experience*. Report for the DAC (Development Assistance Committee) Expert Group on Aid Evaluation, OECD (Organization for Economic Cooperation and Development). Copenhagen, Denmark : DANIDA (Danish International Development Assistance). Retrieved from <http://www.oecd.org/dataoecd/20/52/16546669.pdf>
- Schein, E.H. (2009). *Helping: How to offer, give and receive help*. San Francisco: Berrett-Koehler Publishers.
- Schofield, J.W. (2002). Increasing the Generalizability of Qualitative Research. In A.M. Huberman and M.B. Miles (Eds.), *Qualitative Researcher's Companion*. Thousand Oaks, USA: Sage.
- Schröter, D.C. (2008). *Sustainability evaluation: Development and validation of a Checklist*. Doctoral dissertation, Western Michigan University, 2007. *Dissertation Abstracts International*.
- Scriven, M. (2008). *The Evaluation of training: A checklist approach*. Kalamazoo, MI: Western Michigan University.
- Segone, M. (2005). The Nigerien Way of Monitoring and Evaluation Capacity Development. In S. Mathison (Ed.), *Encyclopedia of Evaluation*, 40-41. Thousand Oaks, CA: Sage.
- Segone, M. (Ed.) (2011). Country-led M&E systems: Better evidence, better policies, better development results. New York, NY: UNICEF Retrieved from <http://www.ceecis.org/remf/Country-ledMEsystems.pdf>
- Sen, A. (1999). *Development as freedom*. Oxford University Press.
- Senge, O. (1990). *The fifth discipline. The art and practice of the learning organization*. New York, NY: Doubleday.
- Serrat, O. (2008). *Conducting after-action reviews and retrospects*. Manila: Asian Development Bank.
- SIDA (2007). *Looking back. Move Forward. Sida evaluation manual (2<sup>nd</sup> revised edition)*. Stockholm: SIDA. Retrieved from [http://www.sida.se/PageFiles/79951/SIDA3753en\\_Looking\\_back.pdf](http://www.sida.se/PageFiles/79951/SIDA3753en_Looking_back.pdf)
- Smith, M. K. (2001). Chris Argyris: Theories of action, double-loop learning and organizational learning. In *The encyclopedia of informal education*. Retrieved from [www.infed.org/thinkers/argyris.htm](http://www.infed.org/thinkers/argyris.htm).
- Smits, P. (2011). Evaluation-capacity building: The three sides of the coin. *The Canadian Journal of Program Evaluation*, 26(1), 75.
- Snowden, D., & Boone, M. (2007). A Leader's framework for decision-making. *Harvard Business Review*, 11, 69-76.

- Solomon, M. J., & Chowdhury, A. M. (2002). Knowledge to action: Evaluation for learning in a multi-organisational global partnership. *Development in Practice*, 12, 346–354.
- Stake, R.E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Stevenson, J. F., Florin, P., Mills, D. S., & Andrade, M. (2002). Building evaluation capacity in human service organizations: A case study. *Evaluation and Program Planning*, 25, 233-243.
- Stockdill, S. H., Baizerman, M., & Compton, D. W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, 93, 7-25. San Francisco: Jossey-Bass.
- Strauss, A., & Corbin (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.
- Stufflebeam, D.L. (2002). *Institutionalization Evaluation Checklist*. Kalamazoo: Western Michigan University.
- Taschereau, S., (1998). *Evaluating the impact of training and institutional development programs: A Collaborative approach*. Washington DC: World Bank.
- Taut, S. (2007). Studying evaluation capacity building in a large international development organization. *American Journal of Evaluation*, 28, 45-59.
- Taylor, P. and P. Clarke. 2008. *Capacity for a Change*. Brighton: Institute of Development Studies.
- Tembo , F. (2008). *Study on capacity development support initiatives and patterns: LCDF research and development phase*. Report for SNV Netherland Development Organization. London: ODI.
- Terry, W., Samset, K., (2010). Issues in front-end decision making on projects. *Performance Management Journal* 41, 38–49.
- Toulemonde, J. (1995). The Emergence of an evaluation profession in european countries: The case of structural policies. *Knowledge and Policy*, 8, 43-54.
- Toulemonde, J. & Rieper, O. (Eds.) (1997). *Politics and practices of inter-governmental evaluation*. New Brunswick, NJ: Transactions Publishers.
- Trochim, W. (1991, June). *Developing an Evaluation Culture in International Agriculture Research*. Invited address presented at the Cornell Institute on International Food, Agriculture and Development's (CIIFAD) workshop on the Assessment of International Agricultural Research Impact for Sustainable development, Cornell University, Ithaca NY.

- Trochim, W. (2006). *An Evaluation Culture*. Web Center for Social Research Methods. <http://www.socialresearchmethods.net/kb/evalcult.php>
- Ubels, J. (2010). Stimulating the provision of local capacity development support. In J. Ubels, Na-A. Acquaye-Badoo, & A. Fowler, (Eds.), *Capacity Development in Practice* (pp. 307-322). Earthscan: London.
- United Nations (2004). *Needs Assessment Report on the Millennium Development Goals*. [http://www.unmillenniumproject.org/documents/mp\\_ccspaper\\_jan1704.pdf](http://www.unmillenniumproject.org/documents/mp_ccspaper_jan1704.pdf)
- United Nations Development Programme (UNDP) (1998). Capacity Assessment and Development in a Systems and Strategic Management Context', *Technical Advisory Paper No. 3*, Management Development and Governance Division, Bureau for Policy Development, January, 1998.
- UNDP (2006). *Capacity development practice note*. New York: UNDP.
- UNDP (2006). *Capacity development strategy for knowledge services and learning*. Learning Resource Center. New York, NY: UNDP.
- UNDP (2007). *Evaluation of results-based management in UNDP*. New York, NY: UNDP.
- UNDP (2009). *Capacity development: A UNDP Primer*. New York, NY: UNDP.
- UNDP (2010). *Defining and measuring capacity development results*. New York, NY: UNDP.
- UNDP (2011). Human Development Index. New York, NY: UNDP.
- UNEG (2010). *Results-Based management handbook*. New York, NY: UNEG.
- UNEF (2012). National evaluation capacity development. Practical tips on how to strengthen national evaluation systems.hoo  
Retrieved from [http://www.unevaluation.org/necd\\_practicaltips](http://www.unevaluation.org/necd_practicaltips)
- Unido (United Nations Industrial Development Organization) (1990). *Report on the evaluation of UNDP-financed and UNIDO-executed projects of technical assistance in metrology*. Vienna: UNIDO.
- Uphoff, N., Esman, M.J., and Krishna, A., (eds.) (1998). *Reasons for success: Learning from instructive experiences in rural development*. West Hartford, CT: Kumarian Press.
- USAID (2011). *Evaluation Policy*.
- Valery, R., & Shakir, S. (2005). Evaluation capacity building and humanitarian organizations. *Journal of Multidisciplinary Evaluation*, 3, 78-112.
- Van Rizin, G.C. (1995). Cluster analysis as a basis for purposive sampling of projects in case study evaluation. *Evaluation Practice*, 16, 109-119.

- Volkov, B., & King, J., (2007). *Checklist for building evaluation capacity*. Retrieved from [http://www.wmich.edu/evalctr/archive\\_checklists/ecb.pdf](http://www.wmich.edu/evalctr/archive_checklists/ecb.pdf).
- Volkov, B. (2008). A bumpy journey to evaluation capacity: A case study of evaluation capacity building in a private foundation. *The Canadian Journal of Program Evaluation*, 23, 175.
- Vygotsky, L. (1997). *Thought and language*. Cambridge, MA: The MIT Press.
- Watson (2006). Monitoring and evaluation of capacity and capacity development. In *ECDPM Discussion Paper, No. 58B*. Maastricht: European Center for Development Policy Management.
- Willer, B.S., Bartlett, D.P., & Northman, J.E. (1978). Simulation as a method for teaching program evaluation. *Evaluation and Program Planning*, 1, 221-228.
- Williams, H., Harris, R., & Turner-Stokes, L. (2007). Northwick park care needs assessment: Adaptation for inpatient neurological rehabilitation settings. *Journal of Advanced Nursing*, 59, 612-622.
- Waddell, S. (2005). *Societal learning and change: How governments, businesses and civil society are creating solutions to complex multi-stakeholder problems*. Greenleaf Publishing Ltd: Sheffield.
- Waldrop, M.M. (1992). *Complexity: The emerging science at the edge of order and chaos*. London: Viking.
- Watkins, R, Meiers, M., Visser, Y. (2012). *A Guide to Assessing Needs: Essential Tools for Collecting Information, Making Decisions, and Achieving Development Results*. Washington DC: The World Bank.
- Watson, D. (2010). Measuring capacity development. In J. Ubels, Na-A. Acquaye-Badoo, & A. Fowler, (Eds.), *Capacity Development in Practice* (pp. 239-249). Earthscan: London.
- Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. Cambridge University Press.
- Westley, F., Zimmermann, B., & Patton, M. Q. (2006). *Getting to Maybe*. Random House: Toronto.
- Wiesner, E.D. (1993). From macroeconomic correction to public sector reform: The critical role of evaluation. In *World Bank Discussion Paper No. 214*. Washington, DC: The World Bank.
- Wiesner, E.D. (1997). Evaluation Markets and Institutions in the Reform Agenda of Developing Countries. In Chelimsky E., & Shadish W.R. (Eds.), *Evaluation for the 21<sup>st</sup> Century- A Handbook*. Thousand Oaks, CA: Sage .

- Wing, K.T. (2004). Assessing the effectiveness of capacity-building initiatives: Seven issues for the field. *Non-profit and Voluntary Sector Quarterly*, 33, 153-160.
- Wijayatilake, D. (2011). Working toward development results: The case of Sri Lanka. In Rist et al. (Eds.), *Influencing Change: Building Evaluation Capacity to Strengthen Governance* (pp. 13-22). Washington, DC: The World Bank.
- Woodland, R., Hutton, M. (2012). Evaluation organizational collaborations: Suggested entry points and strategies. *American Journal of Evaluation*, 33, 366-383.
- World Bank (1994). *Governance: The World Bank's Experience*. Washington DC: The World Bank.
- World Bank (2005). *Building effective states, forging engaged societies*. Washington, DC: The World Bank.
- World Bank (2010). *Project Appraisal Document (PAD) on a proposed grant in the amount of SDR 19.2 million (US\$ 29.9 million equivalent) to the DRC for an establishing capacity for core public management project*. Washington DC: The World Bank.
- World Bank (2012). *Global monitoring report 2012: Food prices, nutrition, and the Millennium Development Goals*. Washington DC: The World Bank.
- World Bank Development Institute (WBI). 2011. *A review of capacity building results measurement in World Bank projects: The need for a systematic approach*. Washington, D.C.: The World Bank.
- World Bank Independent Evaluation Group (IEG). 2005. *Capacity Building in Africa: An IEG evaluation of World Bank support*. Washington, D.C.: The World Bank.
- World Bank Independent Evaluation Group (IEG). 2006. *Annual Review of Development Effectiveness*. Washington, D.C.: The World Bank.
- World Bank Independent Evaluation Group (IEG). 2008. *Using training to build capacity for development: An IEG evaluation of the World Bank's project-based and WBI training*. Washington, D.C.: The World Bank.
- Wubneh, M (2003). Building capacity in Africa: The impact of institutional, policy, and resource Factors. *African Development Review* 15, 165–198.
- Yachkashi, S. (2010). Lesson from below: Capacity development and communities. In J. Ubels, Na-A. Acquaye-Badoo, & A. Fowler, (Eds.), *Capacity Development in Practice* (pp. 194-207). Earthscan: London.
- Yin, R. K. (2002). *Case study research, design and methods*, 3rd ed. Thousand Oaks, CA: Sage Publications

## **Videos**

Indran Naidoo - Monitoreo y Evaluación en Sudáfrica

<http://www.youtube.com/watch?v=fVhfdbJ1Ovc> (last accessed on July 21 2012)

EvalPartners: <http://www.mymande.org/evalpartners>