

***Evaluation Methodology & Baseline Study of
European Commission
Technical Cooperation support***

Final Report

September 2012

Evaluation for the European Commission





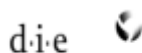
PARTICIP GmbH
Germany



Aide à la Décision
Économique
Belgium



Development
Researchers'
Network
Italy



Deutsches Institut für
Entwicklungspolitik
Germany



European Centre for
Development Policy
Management
Belgium



Overseas
Development Institute,
United Kingdom

A consortium of
Particip-ADE-DRN-DIE-ECDPM-ODI
c/o Particip GmbH, leading
company:

Headquarters

Merzhauser Str. 183
D - 79100 Freiburg / Germany
Phone: +49-761-79074-0
Fax: +49-761-79074-90
INFO@PARTICIP.DE

BELGIUM OFFICE

Avenue des Arts 50 (5th floor)
B-1000 Bruxelles / Belgium
Phone: +32-2-5501160
Fax: +32-2-5501169
INFO@PARTICIP.DE

Framework contract for
**Multi-country thematic and regional/country-level strategy
evaluation studies and synthesis in the area of external co-
operation**

LOT 5:
**Evaluation of EC main policies and strategies in the areas of
external cooperation**

Ref.: EuropeAid/122888/C/SER/Multi
Request for Service: N° 2010/232724/1

Evaluation Methodology & Baseline Study of European Commission Technical Cooperation Support

FINAL REPORT

September 2012

This evaluation is carried out by
DRN

The evaluation was managed by the Evaluation Unit
of DG DEVCO.

The authors accept sole responsibility for this report, drawn up on behalf of the Commission of the European Communities. The report does not necessarily reflect the views of the Commission

STATUS AND VERSIONS OF THE DOCUMENT

Version	Date provided	Status	Validation by	Date of validation
1	03/09/2012	Draft final report	MCB	03/09/2012
2	11/09/2012	Final report		11/09/2012

TABLE OF CONTENTS

TABLE OF CONTENTS	1
1 INTRODUCTION	1
1.1 Objectives and scope of the exercise	1
1.2 Definition of CD	1
1.3 The present report	2
SECTION 1: SUMMARY OF THE LITERATURE ON CD ASSESSMENT IN THE LAST DECADE: KEY EVALUATION ISSUES.....	3
2 THE LITERATURE REVIEW: CONCLUSIONS, LESSONS LEARNED AND PROPOSED COURSES OF ACTION.....	3
2.1 Purpose of the review	3
2.1.1 <i>HOW WAS THE REVIEW DONE?</i>	3
2.1.2 <i>OBSERVATIONS CONCERNING THE SCOPE OF THE EXISTING LITERATURE</i>	4
2.2 Findings concerning CD as an iterative process.....	4
2.2.1 <i>THE 'BLACK BOX' ISSUE</i>	4
2.2.2 <i>PULL VS. PUSH APPROACH</i>	4
2.2.3 <i>BUILT-IN FLEXIBILITY AND INTERACTION WITH THE CONTEXT</i>	5
2.2.4 <i>CAPACITY AND PERFORMANCE</i>	5
2.2.5 <i>FINDINGS CONCERNING CD AS A MULTI-DIMENSIONAL AND POLITICAL CONCEPT</i> 6	6
2.2.6 <i>THE INDIVIDUAL, ORGANISATIONAL AND INSTITUTIONAL LAYERS OF CD.</i>	7
2.3 Findings concerning the definition of the end state of CD effort and how to achieve it.....	7
2.4 Findings concerning the evaluation of CD and other measuring issues	8
2.5 Key synthetic conclusions on the literature review.....	8
3 RECENT WORKS ON CD ASSESSMENT AND DIRECT REFERENCES OF THE PROPOSED METHODOLOGY	9
3.1 ROACH	9
3.2 5Cs	10
3.3 World Bank Capacity Development Results Framework (CDRF)	11
3.4 UNDP: Measuring Capacity	12
3.5 The Proposed Methodology.....	13
3.5.1 <i>TWO WAYS OF ASSESSING A CD PROCESS</i>	13
SECTION 2: PROPOSED METHODOLOGY.....	15
4 THE CD PROCESS.....	15
4.1 The narrative.....	15
4.2 The intervention logic	15
4.3 Level 1 of the IL: enabling factors and CD inputs	19
4.4 Level 2 of the IL: capacity outputs	20
4.5 Level 3 of the IL: Capacity Outcomes	21
4.6 The interaction of the key components of the IL.....	23
5 THE KEY STEPS OF THE EVALUATION PROCESS	24
5.1 Preliminary assessments: the OF and the QCa.....	24
5.2 STEP ONE – on the production of capacity outputs	25
5.3 STEP TWO - on the emergence of CD outcomes.....	26
5.4 STEP THREE: Causality links between the CD inputs and the CD outcomes.....	27
5.5 Link between CD evaluation and standard programme evaluation.....	28

5.5.1	<i>CD AND STANDARD EVALUATIONS ARE NOT SUPERIMPOSABLE</i>	28
5.5.2	<i>THE NEED FOR COMPLEMENTARITY</i>	28
5.5.3	<i>CONDITIONS FOR CARRYING OUT A RAPID CD ASSESSMENT</i>	29
SECTION 3: RAPID ASSESSMENT OF CAPACITY DEVELOPMENT (RAC)		30
6	PRESENTATION OF THE RAC	30
6.1	The challenge.....	30
6.2	The Rapid CD Assessment.....	30
6.2.1	<i>DIFFERENT COMPONENTS</i>	30
6.2.2	<i>THE ADMINISTRATION OF THE QUESTIONNAIRE</i>	30
6.2.3	<i>THE CONTENTS OF THE QUESTIONNAIRE</i>	31
6.2.4	<i>PREPARATION AND PROCESSING</i>	33
6.2.5	<i>RESTITUTION</i>	33
6.2.6	<i>SUMMARY OF THE INPUTS AND TIMES NECESSARY FOR A STANDARD RAC</i>	33
7	THE THREE TESTS	34
SECTION 4: ANNEXES		50
ANNEX 1: EQS AND JUDGEMENT CRITERIA		50
ANNEX 2 - RAC TOR		56

LIST OF TABLES:

TABLE 1 COMPARISON BETWEEN THE AREAS COVERED BY THE CURRENT QSG AND THE RAPID AND THOROUGH EVALUATIONS	14
TABLE 2: COMPARISON BETWEEN THE KEY CD OUTCOMES UNDER DIFFERENT CONVERGENT APPROACHES.....	23
TABLE 3: STANDARD EQS FOR THE PRELIMINARY ASSESSMENT.....	25
TABLE 4: STANDARD EQS FOR THE STEP 1 ASSESSMENT	26
TABLE 5: STANDARD EQS FOR EXPECTED AND UNEXPECTED OUTCOMES.....	27
TABLE 6: TECHNICAL INPUTS FOR A STANDARD RAC	34
TABLE 7 : STEP 1: STANDARD EQS AND JUDGEMENT CRITERIA	50
TABLE 8: STEP 2: STANDARD EQS AND JUDGEMENT CRITERIA	53
TABLE 9: STANDARD EQS FOR THE PRELIMINARY ASSESSMENT.....	61
TABLE 10: STANDARD EQS FOR STEP 1	62
TABLE 11: STANDARD EQS FOR STEP 2	63

LIST OF FIGURES:

FIGURE 1 - ROACH APPROACH TO EVALUATION (FROM THE BACKBONE STRATEGY TOOLKIT) – THE SQUARE IS NOT PART OF THE ORIGINAL FIGURE	9
FIGURE 2: OVERVIEW OF THE 5CS MODEL.....	10
FIGURE 3: CD LOGICAL PROCESS ACCORDING TO THE CDRF.....	11
FIGURE 4: CD LOGICAL PROCESS, ACCORDING TO UNDP (MEASURING CAPACITY, 2010)	12
FIGURE 5: PROPOSED STANDARD IL FOR THE EVALUATION OF CD	17
FIGURE 6: NESTED ILS	18
FIGURE 7: STANDARD AND CD CHAINS OF EFFECTS	58
FIGURE 8: PROPOSED STANDARD IL FOR THE EVALUATION OF CAPACITY DEVELOPMENT SUPPORT	60

ACRONYMS

5Cs	5 Capabilities identified by ECDPM/Dutch Cooperation as the CD Outcomes
CD	Capacity Development
CFRD	Capacity Development Results Framework
DEVCO	EC Directorate General for Development Cooperation
EC	European Commission
ECDPM	European Centre for Development Policy Management
EF	Enabling Factors
EQ	Evaluation Question
IDN	Interim Desk Note (of the present evaluation)
IR	Inception Report (of the present evaluation)
JC	Judgement Criteria
JEU	EC DEVCO Joint Evaluation Unit
LFA	Logical Framework Analysis
M&E	Monitoring and Evaluation
MSC	Most Significant Changes
MTEF	Medium Term Expenditure Framework
OF	Opportunity Framework
QC and QCa	Quality Criteria
QSG	EC Quality Service Group
RAC	Rapid CD Assessment
ROA	Rapid Outcome Assessment
ROACH	Result Oriented Approach to Capacity Change
TC	Technical Cooperation
TOR	Terms of Reference
WBI	World Bank Institute

INTRODUCTION

This evaluation is being commissioned within the framework contract for Evaluation of the EC's main policies and strategies which was signed on 10 April 2007 between the EC and a consortium led by Particip (Germany) and composed of ADE-Aide à la Décision Économique (Belgium), DIE-Deutsches Institut für Entwicklungspolitik (Germany), DRN-Development Researchers Network (Italy), ECDPM-European Centre for Development Policy Management (Belgium), and ODI-Overseas Development Institute (United Kingdom). DRN is in charge of the study.

Objectives and scope of the exercise

This evaluation study aims to:

- develop a detailed results-orientated methodology for the assessment of capacity development in the Technical Cooperation (TC) programmes; *and*
- test it in a number of case studies reflecting a variety of EC intervention regions and modalities.

In line with what is specified in the ToR, the scope of the study covers the following four typologies of TC as defined in the EC Backbone Strategy:

- Capacity Development of organisations and individuals;
- Provision of policy and/or expert advice;
- Strengthening of implementation (of services, investments and regulatory activities); *and*
- Preparation/facilitation of EC cooperation (or broader donor cooperation).

Nevertheless, as agreed in the Reference Group meeting held on 9 March, Capacity Development (first type of TC according to the Guidelines n° 3 of March 2009) is the focus of the evaluation, although other typologies are considered when relevant to achievement of capacity development (second and third types of the Guidelines).

Definition of CD

Capacity Development being the focus of the study, it is important to establish from the outset a clear understanding of its definition. The OECD-DAC definitions are used:

- Capacity: understood as the ability of people, organisations and society as a whole to manage their affairs successfully.
- Capacity development: the process by which individuals, groups and organisations, institutions and countries develop, enhance and organise their systems, resources and knowledge; all reflected in their abilities, individually and collectively, to perform functions, solve problems and achieve objectives¹.

Such DAC definitions build on the UNDP's definition "Capacity development [is] the process by which individuals, groups, organisations, institutions and countries develop their abilities, individually and collectively, to perform functions, solve problems and achieve objectives"².

In this report, the word 'capability' is used as well, as synonymous with capacity, but it is never used when referring to the comprehensive dimension of capacity; it is rather applied to specific capacity areas. For

¹ OECD-DAC - Guidelines and Reference Series Applying Strategic Environmental Assessment: Good Practice Guidance for Development Co-operation, OECD, Paris, 2006

² UNDP - Governance for sustainable human development - A UNDP policy document - Glossary of key terms - 1997

instance, the global capacity of an institution is defined as the coherent expression of four fundamental capabilities³.

Describing the Intervention Logic of a capacity development process, the expressions “capacity inputs”, “capacity outputs” and “capacity outcomes” are used to simplify the distinction between the logical components of a capacity process and the standard components used in a development intervention.

The present report

The methodology was drafted and discussed in the summer of 2011. Two modalities have been defined: a thorough evaluation, and a rapid assessment. The rapid assessment modality was tested in three countries at the end of 2011 and at the beginning of 2012, namely Ukraine, Bolivia and Chad.

The present synthesis report includes a summary of the literature review carried out at the start of the work; a presentation of the revised methodology building on experience to date; and validation of the three tests. It is structured in eight sections:

1. Section 1: Summary of the literature review on Capacity Development and key issues for evaluation
2. Section 2: Thorough evaluation methodology
3. Section 3: Rapid Assessment of Capacity Development (RAC)
4. Section 4: Annexes

³ See the 5Cs approach and the Intervention Logic discussed in Sections 2 and 3.

SECTION 1: SUMMARY OF THE LITERATURE ON CD ASSESSMENT IN THE LAST DECADE: KEY EVALUATION ISSUES

THE LITERATURE REVIEW: CONCLUSIONS, LESSONS LEARNED AND PROPOSED COURSES OF ACTION

Purpose of the review

This is an overview of the conclusions and lessons learned from a broad-scoped literature review of technical cooperation (TC). The literature review itself is a much longer analytical document with a comprehensive reference bibliography; it is annexed to the Inception Report. The present document is meant to simplify the reader's task by conveying a structured set of conclusions and lessons gleaned from a more extensive examination of the literature review document.

How was the review done?

While the mandate given to DRN refers to an "Evaluation Methodology and Baseline Study of European Commission Technical Cooperation Support", the literature review focused on the models for, and the evaluation of, TC within its four purposes as defined by the EC, notably for Capacity Development (CD)⁴.

The literature review was based on a wide cross-section of socio-technical domains including development cooperation; business; health; large institutional management (e.g. hospitals and universities) and modern public sector administration. The over two hundred books, articles, reports and other products consulted were consulted in studies undertaken by or for bilateral and multilateral donors; international financial institutions; business schools and journals; institutional science managers; academics and civil society organisations. An important point to consider is that these items had to be publicly available or in the personal library of the researchers and had to be in English, French or Spanish (the vast majority were in English).

The lion's share of the documentation was linked to evaluation or metrics in one way or another, although many texts were considered if they dealt with capacity development but did not specifically cover evaluation. No attempt was made to seek validation of our interpretations of observations, findings or conclusions with the original authors of the documents.

⁴ This focus was agreed to with the EC on March 8, 2010.

A meta-evaluative methodology was developed for the review, with analytical criteria that were blind-tested for consistency. The principle of rich data was applied to reduce the need to take identical observations into account beyond the point where it could be said that the observations were common to most authors. Findings were clustered and statistically analysed using multi-variate analysis software.

Observations concerning the scope of the existing literature

Overall, there was little in the way of consistency in the quality of published research. Most documents build on or refer to a small number of publications (e.g. from DANIDA and DFID) so the same arguments are repeated. The literature is also not generally supported by empirical evidence so that unfortunately much of the documentation is conceptual and opinion-based rather than evidence-based. The vast majority of non-academic texts do not contain the arguments needed to ascertain research reliability or validity.

As noted by the UNDP and the World Bank in key documents on the issue, there is little consensus on concepts or meanings in this domain, and one important document⁵ clearly makes the case that the term “capacity development” has been overused and overextended to a point at which it has lost its analytical utility.

Of particular interest is the obvious lack of documentation published by, or from the perspective of, recipients. This is, in our opinion, a major weakness in the literature, for much of what is written is designed from a north-to-south perspective, with an obvious paradigm that includes underpinning the management and administration of donors.

Findings concerning CD as an iterative process

The ‘black box’ issue

Almost all documents reviewed contain CD models that are rather static, two-dimensional and linear, focusing on details of the direct “outputs” that need to be generated or the narrative description of the “outcomes” of an intervention, rather than dynamic change management models or measurement systems that include the change processes that actually takes place within the beneficiary (induced results or internalised processes). A simple analogy is the “systems” view where inputs enter a “black box” to emerge as “outputs”. In this case, the outputs enter a black box to emerge hopefully as outcomes. What happens inside the box is not relevant. This simple model was replaced long ago in management science thinking (e.g. in business, large institutions, public sector administrations) because what is important is not the input, output or outcome but the ability of the black box to evolve and change to produce what is needed on a sustainable basis, through an internalisation process that renders the capacity relevant and “owned” (this is obviously a simplified definition).

Pull vs. push approach

Within the literature, the “typical” vision of TC is to see CD within a “functional rational” rather than “political motivational” paradigm. In practice this means that TC is generally used to generate systems, processes and structures (through various strategies including training and mentoring) as part of a CD-push strategy, rather than the organisational and institutional dynamics, motivations, political support and other factors that are seen as prerequisites for change (i.e. CD-pull). The CD concepts found in the literature are therefore relatively fixed in terms of the requisite “steps” prescribed for success. These steps create a hierarchy and a series of products. The literature analysis clearly points to a key influencing factor for this state of affairs: this is what donors need for their project cycles, contract management, and so forth. Evaluations done by major donors support

⁵ A DFID-sponsored evaluation of TC notes that “there is a lack of common approach and terminology to CD”.

See Oxford Policy Management, “DFID Synthesis Report EV 667- Developing Capacity? An evaluation of DFID-funded Technical Cooperation for Economic Management in Sub-Saharan Africa”, 2006. See also Mizrahi, Y., “Capacity Enhancement Indicators-Review of the Literature”, World Bank Institute Working Paper, Stock number 37232, 2004

this analysis and report that because of this and other vectors, the overall scorecard for the success of technical cooperation within a CD context has been very poor for decades⁶.

Built-in flexibility and interaction with the context

Some documents refer to the need to contextualise to local conditions but this aspect is not often further developed even if practically every evaluation report will point to the absolute need to contextualise. In contrast, a relatively constant approach in the non-developmental literature is to position CD (and change generally) as a long-term iterative process, with CD end-products (in terms of capability) defined and with built-in flexibility (and empowerment) in the administrative and management processes. The development literature is instead relatively weak when it comes to describing how to implement this more flexible and adaptive model (including the use of external vectors such as TC)⁷.

Basically, this more current vision (including flexibility, adaptation, emphasis on sustainable performance, etc.) is an open field of research: while it reflects the aims of the Paris, Accra, Busan and other international agreements, the development praxis finds it very difficult to implement, in part because current practice does not take account of the accountabilities of each partner. In fact the literature is relatively silent on how to contextualise models or manage complex processes within a donor-beneficiary relationship that does not have downstream accountability consequences (the literature on governance and accountability is strong on this point).

Capacity and performance

Whenever the literature notes that complexity is a vector that needs to be managed, the documents agree that managing CD as an iterative process is radically different from past TC or CD practice (incl. design, inputs, monitoring, relationships, evaluation etc.). The literature makes it clear that the intervention design carried out for CD interventions and (the rigid and linear) management systems of donors are not geared to the process or systems view needed for CD analysis. A logical consequence of the foregoing is that project management flexibility is, and has been, a real requirement for CD and TC, along with ongoing management of uncertainty and risk; the ability to easily engage in resource mobilisation or demobilisation, motivation, ownership and accountability on the part of recipients and a focus on managing the change process, rather than in the production of pre-defined outcomes. As noted in non-development literature, “performance is not a proxy for capacity”. In other terms, evaluation should measure capacity and not performance. These factors are cornerstones of the non-developmental literature but, according to many evaluation reports studied in this mandate, donors have not set up their procedures and systems with the eventual success of CD in mind. A case in point is that the literature clearly points to monitoring processes that are weak and disjointed, with managers having few levels of empowerment to change and adjust as needed.

Findings concerning CD as a multi-dimensional and political concept

As noted previously, most of the literature focuses on functional “push” rather than political “pull”. Although Grindle and Hildebrand⁸ had provided a comprehensive analysis of the need to include

⁶ Many references support this conclusion. These can be found in the longer text of this literature review.

While almost all experiences with CD are identified as being inefficient in these texts, the reader is advised to remember that CD is contextual and there are a great many published cases where the CD has been successful and sustainable. The World Bank’s “guidance Note for Project Management – Strengthening Institutional capacity during Project Implementation” October 2005, clearly and unambiguously notes that Project Implementation Units have “failed dismally” in terms of any long-term impact on capacity building”. PIU’s and TC in traditional forms of CD have been identified as a serious problem by the World Bank’s Operations Evaluation Department.

⁷ For a more comprehensive comparison between the development and the business sector approaches to CD, refer to the proceedings of the EC’s July 4 Conference on CD held in Brussels in 2011.

⁸ Grindle, M.S. and Hildebrand, M.E., “Building Sustainable Capacity in the public sector: what can be done?” Journal of Public Administration and Development, Vol 15, 1995. This publication forms a key pillar of the World Bank Institute’s 2004 paper on capacity enhancement indicators.

individual, organisational and institutional analytical levels in any CD-based intervention as early as the mid-1990s, the role of motivation and political support was not re-examined in detail until fairly recently. When comparing development praxis and non-development praxis, the authors of this report found that recent publications in the non-developmental domains are more likely to recognise that CD is accelerated or enabled by political pull enhanced by perceived positive motivation, political timing and the support of external lobbies. The above-noted paper describes a substantial analytical framework for assessing capacity and capacity enhancement that is built upon a model that disaggregates capacity for each of the three levels of analysis. They state that although training individuals, transferring capacity and developing more efficient systems is important, capacity depends more on improved management that is able to ensure the integration (and internalisation) of new abilities, skills and knowledge into everyday work. Part of their work reflects Nonaka's and Takeuchi's⁹ model of the transfer of explicit knowledge from an individual (it could be through TC) to tacit knowledge in an organisation (group)¹⁰. These authors repeatedly refer to the absolute need to provide a motivating environment in which development of capacity is encouraged and new capacity is constantly used to "push the envelope". In the specific case of CD within a development context, the literature is quite unanimous on one point: the motivational leadership is - or is not - provided from the highest levels of government. Evaluation reports indicate "mostly not".

As noted in a UNDP research paper published in 2002¹¹, most references to the underlying cause of poor results of CD in development contexts cite weak "ownership", "poor political support" or similar terms. Nevertheless, analysis of real experiences covering the contribution of political motivation to good capacity development is rarely expanded on in the development literature, although very recent evaluations on Tunisia and other countries, where strong GBS programmes have been implemented, provide some light on this subject. In fact, there is emerging evidence that an *increased sense of ownership and support for CD* can be the result of the *political motivation* inherent in socially-important objectives negotiated with the EC (e.g. EU enlargement, EPA, trade agreements). The literature clearly shows that there are few successful examples of externally-driven and motivated CD (i.e. by the donors). An important lesson is that "motivation" needs to be internalised and supported, but the literature is weak on the strategies that should be used by donors to provide support for that motivation. The recent Tunisia experience referred to above has given the question a new dimension: how could the donors (in this case the EC) use "POLITICAL MOTIVATION" to reinforce the partnership relationship, thus providing a self-interested objective for developing partners?. This issue could be important in and of itself, but also as a backdrop for evaluation efforts aimed at learning from experience.

Unfortunately, the literature is relatively silent on how to design change (sometimes called transformation) strategies that are based on beneficiary-pull or political-pull strategies; in fact, some documents assume that this is the role of recipient institutions as part of their accountability and do not refer to it at all.

The individual, organisational and institutional layers of CD.

As alluded to previously, few documents studied as part of the literature review treat CD as a "whole systems" or "holistic-complex" issue, or address it from the level of the societal benefits being sought. Somewhat simplistically one could conclude that most of the documents concentrate either on the linear generation of outputs (which can be produced either by the donor-supplied TC or a combination of "TC and recipient"), or on the need to clearly identify which "outcomes" need to be

⁹ Nonaka, I. and Takeuchi, H. (1995) *The Knowledge-Creating Company*, New York: Oxford University Press

¹⁰ The Takeuchi model has, of course, been modified and criticised over time, particularly the concept of "transfer". The Knowledge Management literature makes a clear distinction between "transfer" (i.e. on the part of the owner of knowledge, and "learning" (i.e. on the part of the recipient. A good overview of these conflicting views and the effect of motivation on learning at an organisational level can be found in Tsoukas, H., "Do We Really Understand Knowledge?" in Easterly-Smith, M. A. et al, "Organisational Learning and Knowledge", Blackwell Publishers, 2003

¹¹ Fukoda-Parr, S., Lopes, C. and Malik, K., "Capacity for Development: New Solutions to Old Problems", UNDP-Earthscan Publications, 2002

produced. Either way the focus is on the performance of the donor intervention, and not on the complex dynamics that result in a sustainable new capacity generated. A recent donor-generated report¹² specifically examined the influence of the performance-framework-defining tools it used (e.g. LFA) and found that the recipients and managers in the field did not find them useful, largely because they were designed to help the donors, not the recipients. The literature review showed that not only is a significant part of the donors' project cycle toolkit seen as benefiting the donor rather than the beneficiary, but also that the majority of indicators used in CD were performance indicators for the intervention (external reference), and not capacity-driven (internal reference).

Quantitatively speaking, the development literature (i.e. the largest number of documents) was focused on the development of the capacity of the individual, while temporally (i.e. what has been most recently written) many documents tended to expand the individual focus to the organisation. But recent publications in the management and sector-institutional domains (e.g. education, health) are now based on strategies that reflect the belief that the institutional context is as relevant. The political context, support from potential beneficiaries, ties between the motivation of the decision-maker and the changes sought, and so forth, are often key to the ability of the organisation to use the new capacity it has acquired (through empowerment and delegation). The authors of the literature review did not find any example of a text that expanded on the interplay between these three levels. Although it would have to be highly context-specific, such a treatise would be a very useful complement to the existing literature. In any case a useful lesson learned is that it is important not only to define and evaluate organisational and sector outcomes, but institutional outcomes too.

Findings concerning the definition of the end state of CD effort and how to achieve it

While every management science text dealing with change management will go to considerable length to speak of the complexity of change and its management, creating models that are iterative, with second-order feedback loops and redundancy, the development literature is weak on this important point. The literature review found that comprehensive capacity assessment guides were only a very recent phenomenon and that they focus the assessment in the hands of external bodies (often NGOs or consultants) rather than the "recipient". In fact, except for twinning arrangements, needs assessments and response strategies are generally guided by donors, not beneficiaries. The majority of CD-related TORs examined by the authors did not contain room for significant "joint" assessments of needs or intervention design with beneficiaries. The authors have formed hypotheses as to why this happens (including the nature of contracting for "external expertise").

Interestingly, the intervention documents reviewed consistently assumed that the most appropriate response for a CD requirement was TC. Even capacity assessment guides did not dwell on the need to consider other options, nor did they describe how to compare options. At the limit, it would have been interesting to study a document that proposed a model for comparing different use-of-TC options, but none was found. Thus a comparison of the effectiveness of various TC options was not addressed in the literature we studied. A notable finding was that the literature generally assumed that TC would provide what is missing (most often in terms of knowledge), rather than change what is already there.

Findings concerning the evaluation of CD and other measuring issues

The following points are worth noting:

- A number of documents propose capacity indicators. Most are general (in the sense of being applicable to just about any CD intervention). They very much focus on the front end of CD processes and need to be contextualised to be useful in specific cases.
- Measurement models and indicators tend to be in terms of increased performance at the individual and organisational levels, not increased capacity. Yermile Mizrahi's seminal paper for

¹² Sida

the World bank Institute¹³ *inter alia* made two points very forcefully: first, that capacity enhancement needs an environment where “recipients” are able to “use” acquired knowledge; and second, that performance indicators cannot be substituted for capacity enhancement indicators. Capacity and performance are not synonymous.

- TC is not discussed with reference to its role within the specific change strategy.
- Baselines are often mentioned as necessary; however, they are rarely described or available to evaluators.
- Monitoring and Evaluation has tended to concentrate on the accountability needs of donors; results in the literature indicate that most M&E systems are disconnected from local learning and knowledge management systems.
- Evaluation tends to be *ex ante* or *ex post*. If CD is a process, then the managers of that change or transition need to rely on ongoing evaluation.
- Evaluation models do not include the need to seek evidence of the influence of exogenous forces acting on the capacity development “system” (i.e. from political, motivation, organisation, culture, etc.).
- The evaluation literature generally does not cover the influence of “participatory” or “ownership” vectors on capacity change.
- Even if Intervention Logic diagrams clearly identify the causality relationship between capacity development and downstream impacts (specific or intermediate), the literature we examined (i.e. final reports of country-level programme evaluations) showed that most often the programme design did not include creation of databases and monitoring systems. This weakness effectively constrains the downstream measurement of CD’s “contribution” to societal-level impacts.

Key synthetic conclusions on the literature review

To conclude this short summary of the literature review, some key points are considered particularly important for constructing the proposed approach to CD assessment:

- *Pull versus Push* approach to CD, meaning taking into account the factors that drive (*pull*) the endogenous processes and how support programmes can adjust to such factors (instead of just trying to *push* the processes).
- *Capacity versus Performance*, distinguishing assessment of capacity from a judgement on the effectiveness of a support programme and on its degree of compliance with the established performance indicators.
- *Knowledge sharing* and learning processes: for capacity development the learning process for acquiring new knowledge and skills should be considered at least as important as the knowledge and skills themselves.
- *What to evaluate and how*. Here there are two main issues to address: [a] how to distinguish CD evaluation from the standard support programme evaluation; and [b] how to integrate CD assessment into the current project cycle management, while limiting its burden and pervasiveness.

RECENT WORKS ON CD ASSESSMENT AND DIRECT REFERENCES OF THE PROPOSED METHODOLOGY

In recent years the main international agencies have accelerated the search for new approaches to identifying and evaluating capacity development. The present study starts from the achievements of

¹³ Mizrahi, Y., “Capacity Enhancement Indicators-Review of the Literature”, World bank institute Papers , 2004

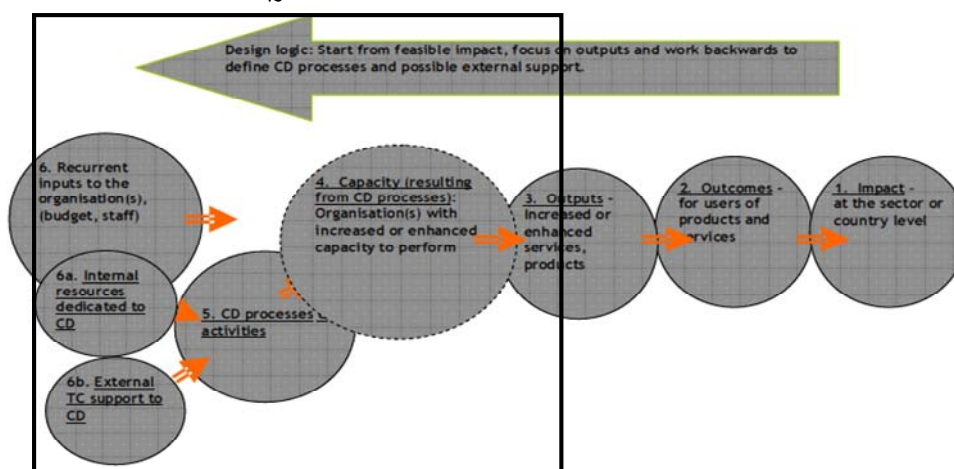
the European Commission's Backbone strategy for Technical Cooperation and builds on its experience. Other important references, however, include: the work of the Netherlands evaluation unit (IOB), which has tested the 5C approach to Capacity Development; and that of the World Bank, which has established the Capacity Development Results Framework and tested new evaluation tools; and that of UNDP, which has further developed its approach for capacity measuring.

ROACH

The Backbone Strategy has supported the ROACH approach to the assessment of capacity development. The present study has developed a detailed analysis of the applications of the ROACH approach¹⁴. It has concluded that this approach highlights well the capacity development process (Figure 1) implicit in the intervention logic of a TC or other support intervention, which aims to transform a set of inputs into a set of development outputs, outcomes and impacts. It does not, however, provide specific instruments for identifying whether or not a capacity development process has taken place successfully, aside from the standard performance assessments. In other words, ROACH does not allow an assessment of the capacity development process with its own specific indicators *vis-à-vis* the output and outcome indicators of the standard support programme. In addition, ROACH – while considering a number of enabling factors that favour the capacity development processes, namely the Quality Criteria – does not consider any “pull” factor.

The CD evaluation methodology developed here builds on the key achievements of the ROACH system, that is to say recognition of the CD process necessary to attain sustainable results. However it integrates and revises the approach, to allow specific assessment of the CD process and to take into account the role of the context-related driving factors in the determination of the CD processes.

**FIGURE 1 – ROACH APPROACH TO EVALUATION (FROM THE BACKBONE STRATEGY TOOLKIT) –
 THE SQUARE IS NOT PART OF THE ORIGINAL FIGURE**



5Cs

Other recent methodological works on CD, already mentioned, respond in different ways to the need for assessing or evaluating the CD process *as such*, and not (or not only) in relation to the development results generated in the framework of a given support intervention. This implies that they seek specific capacity indicators and possible logical paths for CD assessment. The basic idea is that the CD process goes beyond the scope of a given support programme and creates capacities that change the behaviour of beneficiary institutions and are instrumental in the accomplishment of their own missions.

¹⁴ For an in-depth assessment of the ROACH, see the Inception Report of this Study, 08/2010, pages 48-52.

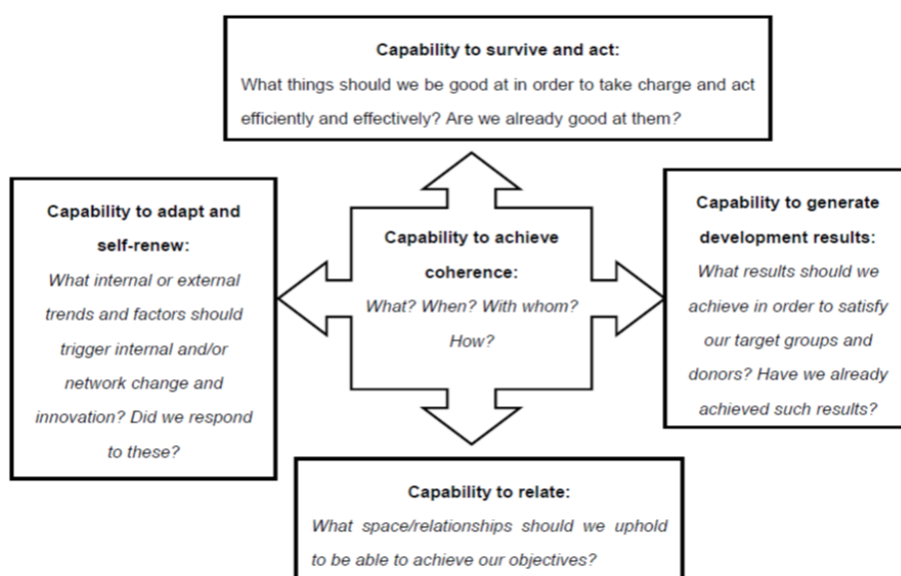
Consideration of the CD process *as such* is particularly important in the 5Cs approach, developed by the ECDPM and supported by the Netherlands IOB¹⁵. The present study, as explained in the following chapters, adopts part of the 5Cs methodology and aims at identifying a fully compatible approach.

Figure 2 indicates how the 5Cs are formulated so as to be individually and collectively linked, since it is the combination and balance of capabilities that defines the capacity of an organisation or system to create value for others. All five capabilities are therefore necessary, yet none is sufficient in itself to create capacity. The model focuses on the endogenous capacity-change processes and insists on coherence between the different capabilities, while not stressing a specific intervention logic with possible intermediate steps (i.e. outputs/outcomes/impacts). The approach is based on “behavioural adaptation to changing environments and conditions”.

It should be noted that evaluating behavioural change is never easy. The Netherlands’ recent experience in the evaluation of CD (using the 5Cs model as part of its methodology) was built around a rather lengthy and complex evaluation process that was difficult to submit to quality control: “transference”. Respondents were asked to speak about their perceptions of change without having any pre-arranged reference point or model construct to refer to. The evaluations then translated the responses into the five capabilities.

FIGURE 2: OVERVIEW OF THE 5CS MODEL

A balanced approach to monitoring and evaluating capacity: assessing change in five core capabilities that affect capacity and performance



World Bank Capacity Development Results Framework (CDRF)

Another relevant reference which deserves particular attention is the World Bank Capacity Development Results Framework (CDRF – 2009). The Framework establishes an approach to orientating the Bank’s programmes towards an improvement in their CD content, with a view to enhancing achievement of the development goals. “The central thesis of the CDRF is that through the acquisition of new knowledge and information – that is, through learning – agents of change can enhance the conduciveness of the socio-political environment, the efficiency of policy instruments, and the effectiveness of organizational arrangements and so contribute to the achievement of

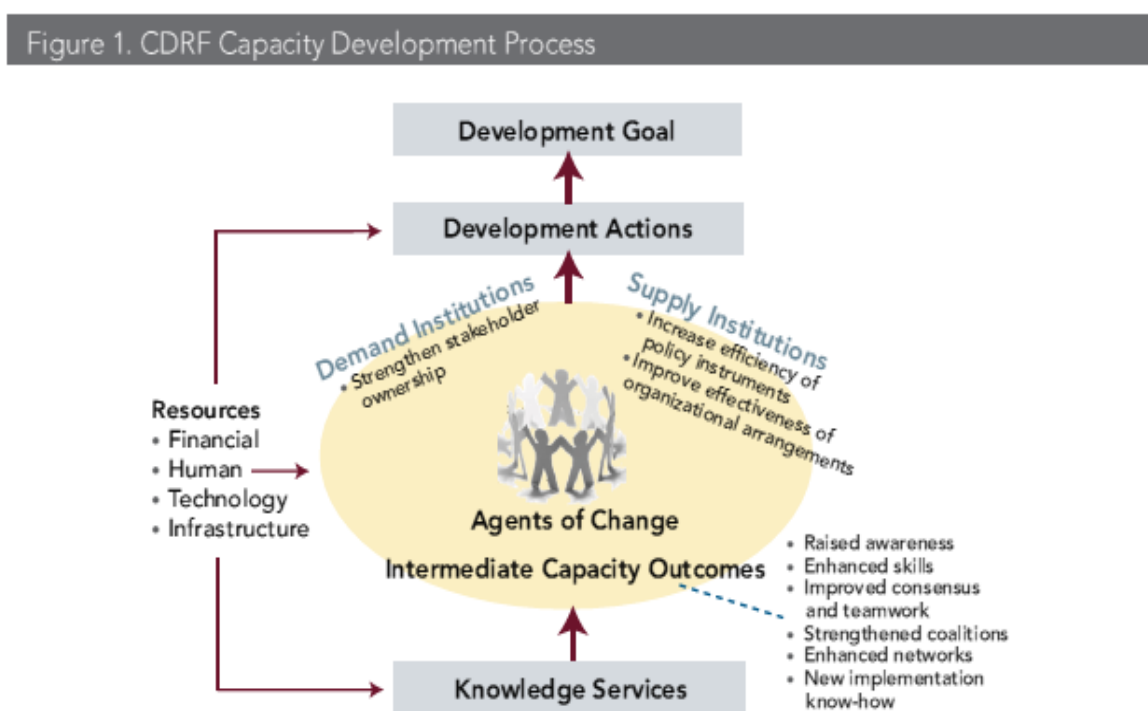
¹⁵ Engel, P., Keijzer, N., Land, T. 2007. “A balanced approach to monitoring and evaluating capacity and performance: A proposal for a framework” (ECDPM Discussion Paper No. 58E). Maastricht: ECDPM. See also: ECDPM, 2011, “Bringing the invisible into perspective”: reference document for using the 5Cs framework to plan, monitor and evaluate capacity and results of capacity development processes.

development goals”¹⁶. The three factors mentioned are regarded as the CD determining factors. An evaluation of the CD process should consider the learning process put in place with the aim of enhancing the three CD determining factors. The Framework establishes a number of detailed indicators for assessing the three factors and tailoring the learning activities of the relevant programmes.

The Framework not only provides the theoretical and methodological instruments for assessing the capacity process *as such*, but also imparts strong importance to the *pull* factors, namely the socio-political framework, and other enabling conditions.

Recently a simplified methodology for the assessment of the CD process, in accordance with CDRF, has also been proposed by the World Bank Institute¹⁷. The approach adopts an Intervention Logic including CD outputs - or intermediate capacity outcomes - and CD outcomes as shown in Figure 3.

FIGURE 3: CD LOGICAL PROCESS ACCORDING TO THE CDRF



UNDP: Measuring Capacity

In its Measuring Capacity of June 2010, UNDP proposes a comprehensive approach to assessing and measuring CD through the identification of specific CD outputs and outcomes that ensure the achievement of the national development goals (Figure 4). The latter are regarded as the impact of the Capacity Development process. UNDP has been particularly advanced in such matters since 1998, having supported the idea of CD as a process.

Now it proposes a very comprehensive approach based on the key assumption that the overall development process depends upon Capacity Development. Again a disaggregation of specific

¹⁶ Samuel Otoo, Natalia Agapitova and Joy Behrens, The Capacity Development Results Framework: A strategic and results-oriented approach to learning for capacity development, World Bank, June 2009, page 15.

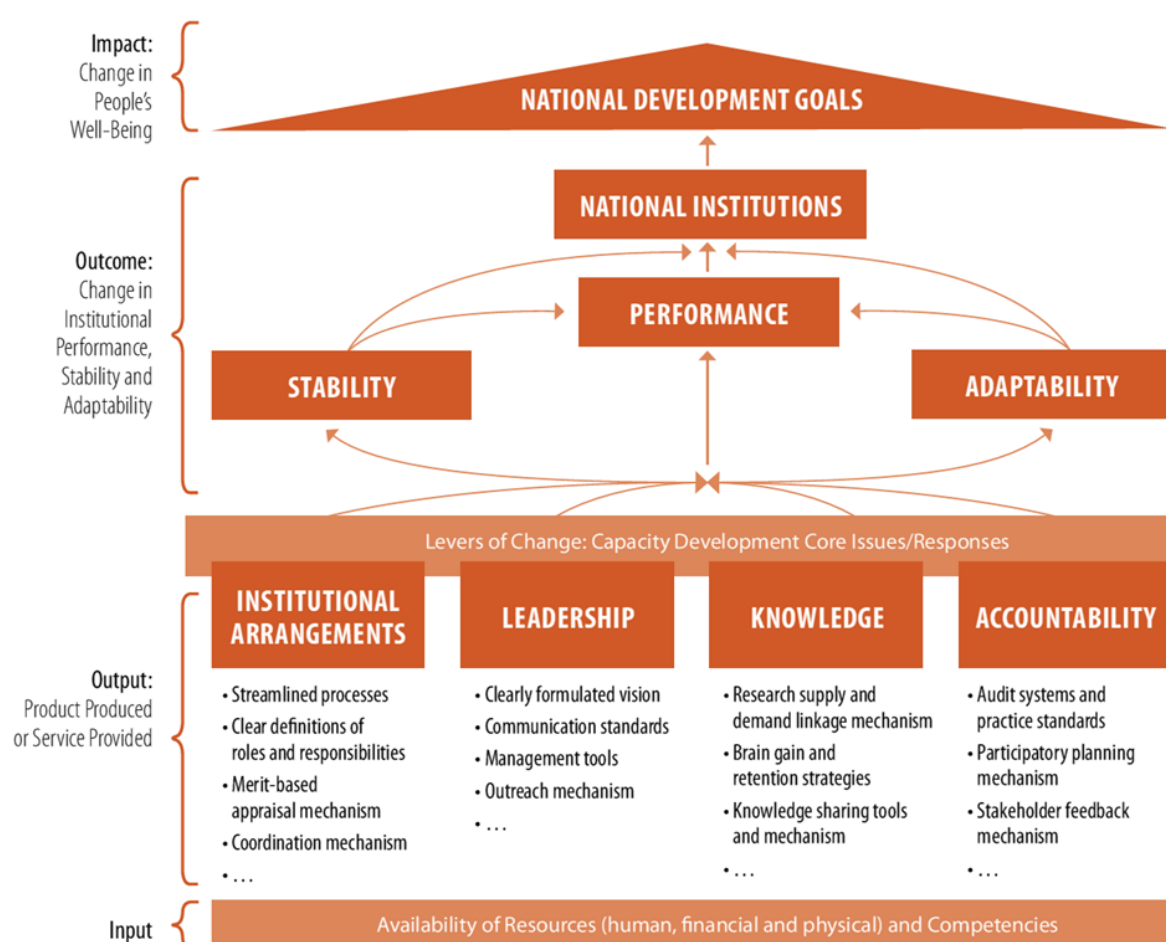
¹⁷ World Bank Institute, Reviewing Project Results Retrospectively Using a Results-Focused Approach to Capacity Development. WBI June 2011

outcomes and outputs is proposed: the CD **outputs** are disaggregated by *institutional arrangement, leadership, knowledge and accountability*, while the **outcomes** are identified as *performance, stability and adaptability*. The theory of change is that the outputs, duly developed and integrated, will generate the outcomes in terms of comprehensive new capacities of the national institutions, which will become globally stronger and empowered to achieve the national development goals.

The UNDP model preceded the WBI model by a few months. The two are very similar, although the WBI model emphasises the distinction between demand and supply institutions and proposes a slightly different breakdown and definition of what the UNDP model defines as outputs and outcomes.

None of the assessment models examined so far gives any specific consideration to the external driving factors, although they consider that the external environment has an importance of its own, as in most evaluations

FIGURE 4: CD LOGICAL PROCESS, ACCORDING TO UNDP (MEASURING CAPACITY, 2010)



The Proposed Methodology

The proposed methodology will build on the recent evolution of the approach to CD evaluation as outlined in the paragraphs above. The key points supporting the theory of change, on which the proposed methodology relies, are highlighted below:

1. The Capacity Development process of a beneficiary institution (or institutional system) cannot be assessed through the performance indicators of a single specific support programme.

2. Capacity Development is an endogenous process of empowerment, driven by internal and external factors, which enables the institution or system involved to fulfil its own mission.
3. One main point of the proposed methodology is the emphasis on the driving role of the political and economic environment in which the beneficiary institution or system is embedded, which determines its opportunity framework.
4. Another main point of the proposed methodology is the attempt to highlight a relationship between the acquisition by the institution or system of a series of specific individual or organisational capabilities, and the transformation and mainstreaming of such capabilities into a new overall capacity to fulfil its own mission. The identification of specific indicators for both the specific capabilities and the overall capacity is essential for grasping and assessing such a relationship.
5. The opportunity framework and the quality of the learning process act as the catalysers of the Capacity Development process, including the acquisition of the specific capabilities and their transformation or mainstreaming into overall capacity.

Such points and theoretical implications are developed in the next chapters.

Two ways of assessing a CD process

The CD process implicit in any support programme may be assessed in two ways.

- A **quick assessment** based on the methodology can be partly incorporated into the EC Quality Support Group process with minor changes and additions to the present QA and monitoring procedures that use the TC Quality Criteria. The present practices should embody a wider-scope criteria grid as well as a small number of related questions including the two new context-related enabling factors (the Opportunity Framework, as shown in the description of the IL – see also the EQs). This will entail slight changes to the QCa, which should be better focused on CD. In addition the EUDs, through specialised and specially trained consultants recruited locally, should organise a stakeholder consultation to carry out a Rapid Assessment of CD (RAC)¹⁸, the specifications of which have been tested in the three field assessments mentioned. The RAC should identify and map both expected and unexpected outcomes.

The RAC is a relatively short exercise that should be easily carried out on all cooperation initiatives designed to have a significant effect on CD, including actions not classified as TC (e.g. NGO grants, Budget Support, etc.). The RAC – when available – may be used as a key information source in standard programme evaluations, including country and thematic evaluations.

- A **thorough evaluation** of the CD process should be carried out for those actions expected to have a strong effect in terms of CD, where full involvement of the evaluated institution is ensured. In theory one might even suppose that some actions could have only a CD outcome, without pre-determined induced outputs, that is to say with a set of unexpected induced outputs and development outcomes. This already happens with some innovative decentralised cooperation programmes (e.g. EUROSOCIAL-1), of which the expected outcome is a strengthening of the autonomous capacities of the institutions involved without reference to specific technical or policy achievements.

Owing to the invasive dimension of this type of evaluation in relation to the institutions involved, a full commitment from the development partner is required. Such commitment will be achieved only if the institution has a specific interest in the evaluation, so as to facilitate a real joint exercise.

Indeed, a thorough evaluation should acquire and analyse the documentation relating to the internal life and the internal products of the institution, instead of limiting its investigation to

¹⁸ A rapid assessment tool has been proposed by Harry Jones, Simon Hearn in an ODI paper (2009) *Rapid Outcome Assessment (ROA)*, which provides a rapid application of the outcome mapping approach, to assess policy changes. Our proposal is focused on CD and is based on a rapid application of the evaluation methodology proposed in the previous section, through the use of participatory techniques, including Most Significant Changes (MSC), as the main research tool.

interviews, group meetings or workshops with the key stakeholders. It might raise concerns and opposition unless the full commitment of internal and external decision-makers is ensured.

The time issue is also particularly important. The possibility of the results of the evaluation being available within a short time is part of the value added of a CD evaluation, which should become an instrument for further institutional debate and strengthening.

When a thorough evaluation is decided upon, ToR would be drawn up based on the EC JEU evaluation guidelines. The standard evaluation approach would be used with its various phases, field missions, Reference Groups and seminars, and the evaluation team would be instructed to use the evaluation model proposed therein.

TABLE 1 COMPARISON BETWEEN THE AREAS COVERED BY THE CURRENT QSG AND THE RAPID AND THOROUGH EVALUATIONS

QSG	RAC	Thorough Evaluation
Quality Criteria Assessment	Assessment of the Quality Criteria, based on EUDs' and other quickly obtainable documentation	Assessment, including additional studies, of the Quality Criteria
	Assessment of the OF, based on EUDs' and other quickly obtainable documentation	Assessment, including additional studies, of the OF
	RAC consultation to assess Capacity Outputs and Outcomes, their causal links and the links with the programme inputs and the enabling factors. The consultation uses participatory methods (including questionnaires, group-coaching, MSC techniques) to acquire and process the relevant information.	Step 1: Capacity Outputs assessment in relation to inputs and context
		Step 2: Capacity Outcomes assessment in relation to outputs and context
		Step 3: Causal links between inputs and Capacity Outcomes and the role of the context

Section 2 presents the Thorough Evaluation.

Section 3 presents the Rapid Assessment of CD (RAC)