

Terms of Reference

Joint Scandinavian evaluation of support to capacity development

1. Introduction

Development assistance has always had the ambition of delivering sustainable results, and, by implication, foster endogenous capacities that eventually would make aid redundant. Skills training and technical assistance delivered inside individual organisations have been among the main inputs expected to create capacities that could deliver sustainable outcomes.

Numerous reviews and evaluations have indicated that expectations did not match reality¹. Attention has also been drawn to the potential negative effects of excessive reliance on technical assistance and training, such as cost, distortions in local labour markets, disruptions in formal hierarchies, weak and twisted accountability mechanisms, and distorted incentives through e.g. salary supplements and workshop allowances.

Even if the term “technical assistance” is still in use, capacity development (CD) is today seen as a much more comprehensive process in theory and development practice. The mainstream view² has been that capacity development is first and foremost an endogenous process where outsiders can at best contribute, but they can normally not claim attribution. The drivers and constraints to capacity development include incentives and performance in the specific context, as well as the interests and priorities of key stakeholders, which shape the arena for support to CD. However, even if this is a dominant message in evaluations as well as donor guidance, it still seems that these insights have not always been transformed into practice.

In parallel with the broadened view on capacity development, donors have over the last decades insisted on results-based approaches, also in the area of CD. Despite the focus on results, it has been difficult to provide hard evidence as to whether capacity development support actually contributes to strengthened endogenous capacities and performance. This also means that it has been difficult to verify the mainstream view that more recent forms of support to capacity development – contextually well aligned, results-oriented approaches – are likely to be more effective.

¹ E.g. Arndt, Channing (2000): “Technical Co-operation”, in *Foreign Aid and Development. Lessons Learnt and Directions for the Future*. Finn Tarp and Peter Hjertholm (eds). London: Routledge.

² See DAC (2006): *The Challenge of Capacity Development: Working Towards Good Practice*. Paris, OECD. See also the five “Perspective notes on Capacity Development” prepared by the OECD/DAC ahead of the 2011 Busan High-Level Forum (<http://www.oecd.org/dac/governance-development/capacitydevelopmentourkeypublicationsanddocuments.htm>) as well as the “Cairo-consensus on Capacity Development” from March 2011 (available on same webpage).

Over the last decade, we have also seen emerging interest in interventions that go beyond the actual institutions expected to improve their capacity. The assumption is that the dominant approach of working from the inside in public sector organisations (“supply side focus”) may be insufficient or even ineffective if not also working on political, legal and other external factors, as well as strengthening the demand for accountability from citizens. This “demand side approach”, while heralded in theory, has not yet demonstrated its effectiveness through evidence-based evaluations.

Another key issue in capacity development is the question of who sets the priorities with regard to the more specific rationale and objectives for capacity development. In line with the Paris agenda, one might expect that the centre of attention would be on strengthening general capacities within given sectors. Nonetheless, efforts to support capacity development may target the capacity of institutions to improve delivery of aid-financed services specifically, or may address aspects of capacity deemed to be of particular importance to donor priorities, rather than aiming at more general capacity development. A distinction between ‘aid effectiveness’ and ‘development effectiveness’ may be relevant here.³

Throughout these different developments in the theory and practice of capacity development, an underlying key issue has been the broad range of relations between donors and partners. This touches issues such as characteristics of the relationship between partners with regard to trust, mutual respect and accountability, the legitimacy of donor interventions, the actual roles each partner play and the incentives for both partners to pay attention to the often delicate and cumbersome processes of change, and the ‘ownership’ by each partner to the processes and results.

This Joint Scandinavian Evaluation aims to cast light on the issues above. It will consist of three separate, but closely coordinated evaluations covering support to capacity development by Denmark, Norway and Sweden, respectively. These Terms of Reference lays out the evaluation commissioned by Danida and covers Danida’s support to capacity development. Similar Terms of References, with some agency-specific amendments, have been developed for parallel evaluations commissioned by Sida and Norad. The three evaluations will respond to the same questions, while each agency may prioritise to look into additional areas of particular high interest. The findings across the three evaluations will be presented in a Synthesis Report based on the individual agency reports.

While focus is on the support to CD from the three agencies, the evaluation is based on the recognition that because capacity development is first and foremost an endogenous process, it is not meaningful to look at what the agencies are doing without seeing this in the wider picture of the efforts of the partner institutions and the context within which this takes place. That may point to recommendations about when donor engagement in capacity development in partner

³ Stern, Elliot D. et al: Thematic Study on the Paris Declaration, Aid Effectiveness and Development Effectiveness. Ministry of Foreign Affairs of Denmark, Evaluation of the Paris Declaration 2008.

institutions is appropriate and legitimate, and under which circumstances donor support to capacity development is likely to be effective.

The field of capacity development is characterised by broadly defined concepts, reflecting the heterogeneity of the field. The OECD/DAC's definition from 2006⁴ will serve this evaluation: "Capacity is understood as the ability of people, organizations and society as a whole to manage their affairs successfully. ... 'Capacity development' is understood as the process whereby people, organizations and society as a whole unleash, strengthen, create, adapt and maintain capacity over time." In this evaluation, the focus will be on capacity development for organisations – acknowledging that both individual and system capacities may be a part of what is required to make an organisation (or a group of organisations) or institution perform better.

As background notes to the evaluation the Scandinavian agencies have commissioned three studies that will inform the evaluation:

- Literature Review for the Joint Evaluation on Capacity Development⁵
- Methodological approaches to evaluate support to capacity development⁶
- Annex I: Approach Paper

The evaluation will be guided by the Approach Paper (Annex I). The Approach Paper expands on the issues mentioned above and lays out an analytical model and generic theory of change behind capacity development support, to enable a shared approach and methodology findings across the three evaluations.

The primary intended users and audience for this evaluation are management and staff within Danida, as well as Sida and Norad. Intended secondary users are other aid agencies, ministries of foreign affairs, and various intermediaries involved in development cooperation including multilateral institutions and governments and institutions in partner countries. Outcomes of the evaluation will also be communicated to the general public and political systems for accountability purposes.

2. Evaluation purpose

The purpose of the evaluation is to improve decision-making and strategy development regarding support to capacity development in developing countries. The evaluation will in particular assess the *relevance* and *effectiveness* of the Scandinavian agencies' support to capacity development, and will address issues of efficiency. It may also generate knowledge about the sustainability and impact of the support to capacity development.

⁴ DAC (2006).

⁵ See: <http://www.sida.se/English/About-us/How-we-operate/Sida-Evaluation/Ongoing-evaluations/Capacity-development/>

⁶ See: <http://www.sida.se/English/About-us/How-we-operate/Sida-Evaluation/Ongoing-evaluations/Capacity-development/>

This purpose has both learning and accountability elements.

With regard to learning, the evaluation aims to produce knowledge that enables policy, strategy and decision makers to design good strategies for support to capacity development and to review, adjust or discard planned and ongoing interventions based on previous experience with support to capacity development.

With regard to accountability, the evaluation aims at assessing results of support to capacity development and to what degree it represents value for money in terms of both relevance, effectiveness and efficiency.

By contributing to a better understanding of how to manage for results in a relevant and adequate manner, the evaluation aims at improving both learning and accountability in future support to capacity development.

3. Focus areas

The evaluation will look particularly at some focus areas seen as critical dimensions of capacity, capacity development and support to capacity development. They are briefly described below, and further explained in the Approach Paper (Annex I):

- i. The relevance and opportunity of a “best fit” approach for CD support well adapted to specific intra- and inter-institutional dynamics and the wider context.
- ii. Within the “best fit” dimension, the appropriateness and the legitimacy of external (donor) involvement in different dimensions of capacity development, and whether some processes may be so complex and demanding that the ability of donors to add value is limited.
- iii. The merits of looking beyond the supply side of public sector institutions to foster broader accountability relations or other types of collaboration with e.g. civil society, private sector, media or oversight institutions.
- iv. How a results-focused approach to aid for capacity development can serve to improve learning and accountability among aid agencies in the future.

4. Scope and delimitations

The evaluation addresses aid that has an *explicit* intention to support institutional capacity development in the recipient country, be it as a primary objective or as integrated components of strategies and programmes having other primary objectives. This may include capacity development pursued with targeted inputs provided to specific institutions as well as interventions addressing factors external to the institution (for instance, by stimulating

accountability via non-governmental institutions) and capacity development expected to happen as a result of the way support is given (i.e. budget support).

The evaluation will focus on support to public sector institutions. Interventions addressing private and non-profit institutions may be included if directly relevant to public sector capacity or if there are other reasons to assume that examining those interventions can shed light on key aspects of support to capacity development (for example, by demonstrating promising practices).

Selection criteria for which interventions to study in-depth will be decided early in the inception phase based on the portfolio screening described in section 6 (approach and methodology) and Annex II.

When assessing results, this evaluation will focus on the achievement of planned outcomes of donor support, as well as to which degree this correlates with actual capacity development in more general terms, acknowledging that the latter depends primarily on other factors than aid.

5. Evaluation questions

The evaluation will be designed to respond to the following questions, based on the study of selected interventions:

- 1) How can a generic theory of change for support to capacity development be formulated that would enhance the effectiveness of support to capacity development?
- 2) What is the relevance of the strategies and initiatives for support to capacity development? E.g. do they primarily aim at improving capacity to manage aid programmes, versus aiming at more general improvement of capacity in a sector or an institution?
- 3) To what degree are the capacities to manage capacity development processes– e.g. change management competencies, incentives, procedures, guidance, management – effectively in place and adequate among the donor agencies and partner institutions?
- 4) How have strategies and interventions been designed to fit with context-specific factors such as specific institutional dynamics or the social, cultural, political and legal environment, and to contribute to influencing factors external to the institution(s), such as demand and accountability mechanisms? To what degree are strategies based on evidence on how support to capacity development has worked elsewhere?
- 5) How do representatives of the partner institutions and/or other stakeholders in partner countries perceive the donors' role in capacity development, and what do they think is the appropriate role of donors in future capacity development?
- 6) How has results-orientation and results-based management approaches been applied in CD support, and how have they contributed to learning and improved effectiveness?
- 7) To what degree have interventions achieved the planned results at outcomes level, and to what degree is there a correlation between the interventions, and observed improvements in capacity of the partner institutions?

- 8) What are the possible unintended effects (positive and negative) of support to capacity development?
- 9) To what degree can one conclude that interventions to support capacity development have been effective and represent good use of resources (value for money), compared to possible other ways of supporting the same sectors or institutions(s)?
- 10) What characterises support to capacity development that is relatively more successful versus strategies and interventions that are relatively less successful?
- 11) Under which circumstances, for which aspects of capacity and for which specific inputs may donor support to capacity development be appropriate and effective? Are there situations where the agencies should refrain from being involved in capacity development, and/or modalities and approaches they should no longer apply?

Danida specific evaluation questions:

- 12) To what degree is Danida following its guidelines to CD support?
- 13) Are the current CD approaches an effective way to reach the poor (directly and/or indirectly)?
- 14) What are the CD lessons learnt which Danida could use to move forward working with new actors in development?
- 15) Given last decade's focus on results-based management, how could Danida work with clearer definitions and reporting on results in CD?

6. Approach and methodology

The nature of the evaluation object poses some challenges with regard to methodology and data issues, including questions around whether certain indicators precisely reflect key aspects of capacity development. There are also limitations to the degree to which changes can be attributed to aid; an enormous heterogeneity of aid supported interventions, as well as heterogeneity of organisations and country contexts.

The evaluation will apply an approach that optimises the likelihood of producing evidence-based assessments and that is realistic given the limitations identified above as well as time and resource constraints. The methodological approach is informed by the methodology study developed for the purpose of this evaluation⁷ and based on the conceptual and analytical models laid out in the attached Approach Paper (Annex I).

The inception phase will include a preliminary screening of a larger sample of capacity development interventions⁸, followed by desk-based study of a smaller sample. This will result in

⁷ See: <http://www.sida.se/English/About-us/How-we-operate/Sida-Evaluation/Ongoing-evaluations/Capacity-development/>

⁸ Danida will provide a list of interventions to be included in the portfolio screening, including a collection of initial documentation to be reviewed.

a standardised set of data collected for each intervention. The aim is both to inform the remaining phases of the evaluation, and to compile data from all three Scandinavian evaluations to enable future statistical analysis beyond the assignment laid out in this Terms of Reference. The details for this phase are described in Annex II.

The main evaluation phase will include three country studies. These will encompass Danida's support to capacity development over a given time period in the three selected countries. Each country visit will comprise about six to nine work weeks combined for all relevant team members⁹. The evaluation team will propose the specific design of the country studies, guided by the Approach Paper (Annex I) and methodology study¹⁰.

Both the inception phase and the main evaluation phase will be coordinated with the other two evaluation teams and the three Scandinavian clients. Danida will have the final word in approving the methodological approach.

When analysing data, the evaluation will apply theory/-ies of change as one analytical approach. The generic analytical model and specific theory of change outlined in the Approach Paper should be used as a starting point unless an alternative proposed by the consultants has been accepted. The theory of change is (as all theories of change) a hypothesis, and the evaluation aims to test to what degree the interventions under evaluation fit with this hypothesis, followed by suggestions for revised or alternative formulations of a theory of change that may serve to explain the findings and provide directions for future CD support.

When assessing results of support to capacity development, focus will be on to what degree programmes achieve their owned planned outcomes, as well as a broader view on to what degree they are likely to have contributed to improved capacity and/or better performance of the institution.

Due to the nature of support to capacity building, where aid interacts with many other internal and external factors that are likely to be stronger determinants for capacity development, in most cases the evaluation will not be able to conclude on attribution. The contribution of aid to observed capacity improvements should be assessed based on the in-depth and country case studies of selected interventions, using theories of change or other analytical approaches.

Capacity can be assessed by looking at organisational capacity parameters (e.g. enhanced systems, processes, skills, management, internal relations etc.) as well as actual performance of the organisation, whether in terms of quality, quantity, cost or relevance or a combination of these. Due to the diversity of the evaluation object, improvements in capacity must primarily be

⁹ Those six to nine weeks will include all work by team members including senior national experts to be recruited after countries have been selected, but excluding junior assistants or other national support.

¹⁰ See: <http://www.sida.se/English/About-us/How-we-operate/Sida-Evaluation/Ongoing-evaluations/Capacity-development/>

measured against improvement in indicators specific to the interventions and institutions, rather than standardised indicators.

The evaluation team may, alternatively, propose an approach that responds to the purpose in this Terms of Reference in other ways than those laid out above and in the Approach Paper (except for the preliminary portfolio screening and review), demonstrating comparable rigor and ability to address the evaluation questions and focus areas. If it does, it should, to the extent feasible, frame its proposal in ways that are compatible with concepts and models of the Approach Paper, to enable coordination and comparison with the evaluations in the other Scandinavian countries.

7. Organisation

The evaluation shall be managed by Danida, which will have the final word in approval of the methodological approach and deliverables. Danida-financed evaluations should be utilisation-focused, meaning that they are designed and implemented along with the intended users and that intended use is at the centre of the evaluation process.

The mechanism for consultation and quality control will be threefold:

- (i) The evaluation Steering Group consisting of representatives from Danida, Norad and Sida. This group is the decision making body in regards to all aspects of the approach and methodology which will cover the joint elements of the evaluation.
- (ii) An Advisory Group composed of representatives from partner countries and donor representatives. The role of the group is to guide and provide feedback to the three parallel evaluations during the inception phase, draft and final reports.
- (iii) A Danida Reference Group. This group will consist of representatives from Danida with the role to give feedback and advice on the agency specific parts of the evaluation (see the Danida specific evaluation questions, section 5).

Representatives of each evaluation team will meet with the Steering Group shortly after contract signing, at the end of the inception phase, and after country visits, at dates and venues to be decided by the Steering Group. The purpose of the meetings are to share findings and ideas and to discuss key issues to lay the foundation for a Steering Group decision on the way forward, and to coordinate the work between evaluation teams¹¹. The communication between the evaluation team and the advisory group will likely be via email. Each team is accountable only to its contracting authority, which will clarify any issues relating to discussions and decisions in the Steering Group and other forums and how to follow-up.

The consultant, within the management framework defined by the respective evaluation department, will be responsible for the implementation of the evaluation in line with the

¹¹ The meetings will be held in the different Scandinavian capitals in turn and the teams should budget for one travel to each of the capitals for these meetings.

principles of independence and impartiality. The consultants shall in their proposal also specify how quality assurance will be handled by them (see chapter 11).

The evaluation will be organised into four work phases; (i) inception phase; (ii) country visits; (iii) analysis and report writing; and (iv) dissemination. The main parts will be carried out over the period October 2014 – June 2015, while dissemination is planned for fall 2015. Each phase is associated with certain deliverables, specified below.

8. Deliverables and time frame

Unless otherwise agreed during the inception phase, the evaluation will involve the following deliverables, including written products as well as presentations and participation in relevant meetings. All reports shall be written in English and adhere to the OECD/DAC quality standards for evaluation.

a) Preliminary portfolio screening note

The team shall deliver a draft, preliminary note from the portfolio screening (Annex II), including identification of samples for the desk-based review and a preliminary indication of countries that seem appropriate for the country studies.

b) Inception report

The team shall deliver an inception report not exceeding 30 pages, excluding annexes, and including, but not necessarily limited to:

- A brief historical background of Danida's work with capacity development and its current approach,
- The results of the portfolio screening and the desk-based review (see Annex II),
- Elaboration on the evaluation approach and evaluation questions and how to respond, including a strategy for all necessary data collection and analysis, and a discussion on limitations,
- Proposed selection and methodological approach for the country studies,
- A detailed work programme,
- A draft Table of Contents for the main evaluation report,
- A draft communication plan.

c) Country studies

Findings and conclusions from the country studies shall be presented separately as stand-alone working papers, not exceeding 10 pages excluding annexes. The main contents shall be discussed at wrap-up meetings in each of the countries visited, then revised and submitted to Danida as draft country reports.

Each country visit shall deliver at least ten JPEG pictures illustrating Danida's support to CD. The pictures may, for example, illustrate stakeholders (including beneficiaries), Danida specific solutions, or the bigger picture.

The team leaders will meet with the three agencies in a joint meeting in a Scandinavian capital city to present and discuss the country reports followed by a discussion on commonalities across the country studies and possible common or joint approaches of relevance to the remaining data collection and analysis. The presentation may include an outreach event to invited participants by the Scandinavian agencies.

d) Main reports

The main report shall synthesise results from the inception phase including the portfolio screening, the desk-based review¹², as well as the country studies. Apart from responding to all parts of this ToR and requirements further detailed during the inception phase, it shall to the greatest possible extent present actionable recommendations.

The report shall not exceed 60 pages excluding annexes, and shall include an executive summary, draft acknowledgment as well as a draft back cover text.

An evaluation brief shall be delivered based on the findings from the final report. The evaluation brief shall follow Danida's guidelines for evaluation briefs.

In addition, the team leader shall contribute to the process of producing a synthesis report for the three parallel evaluations carried out by Danida, Norad and Sida. This will include working in close collaboration with the two other team leaders as well as an assigned consultant responsible to coordinate and finalise the synthesis report. It is anticipated that each team leader must allocate one week of work for the synthesis report.

e) Dissemination of results

The team leaders shall present the final evaluation report and the synthesis reports at a workshop in a European capital city organised by the Steering Group, as well as a workshop in Copenhagen during fall 2015.

¹² Also see Annex II for reporting guidelines for the portfolio screening and desk-based review.

Table 1: Tentative time plan

Time	Activity
Medio September 2014	Signing of contract
Ultimo September	Start-up workshop in a Scandinavian capital to agree on a common way forward as well as the methodology for the joint parts of the evaluation.
15 October	Preliminary portfolio screening note with identification of samples for desk studies and selection of country studies
15 November	Draft inception report
15 December	Final inception report
Primo December	Inception workshop in a Scandinavian capital to conclude on key issues regarding methodology and present initial findings from the portfolio screening.
Medio December 2014 – March 2015	Country visits
20 March, 2015	Draft country working papers
March/April	Workshop to discuss findings from country visits in a Scandinavian capital city.
20 April	Final country working papers
8 May	Draft evaluation report
29 May	Final evaluation report
June	Provision of inputs to evaluation Synthesis report
30 June	Draft synthesis report
30 August	Final synthesis report
Fall 2015	Two dissemination events: (i) in a European capital; and (ii) in Copenhagen.

9. Resources

The Consultant's financial proposal shall include all cost for key personnel fee and staff related expenses as well as project related expenses, and any fixed amounts, which the Consultant shall include for any special purposes for the assignment, e.g. cost of arranging seminars or study tours, funds to be administered by the Consultant, etc.

The maximum budget for the evaluation including provisional sums is 2.5 million DKK. In the price proposal the tenderer shall include a provisional sum of 100.000 DKK for workshops and other dissemination. The evaluated price proposal is the overall price less the provisional sums.

10. Team qualifications

The evaluation is expected to be carried out by a team of minimum three consultants, including consultants who have considerable experience with capacity development issues at the international level and consultants with working experience from country level.

The tender proposals prepared by the tenderer shall include:

- An initial elaboration of the proposed evaluation approach and methodology including approach for data collection and validation,
- Approach and criteria for selection of evaluation themes and priority selection of capacity development funded programmes and projects (the tender proposal shall not include selected programmes and projects to be evaluated; these will be decided in the inception phase through the portfolio screening process and in dialogue with Danida).
- An outline of an initial evaluation framework/matrix and evaluation plan (to be further elaborated in the inception report),
- A draft work plan with relevant process milestones,
- Comments to the terms of reference and evaluation approach paper.

The Consultant should be able to draw on a combination of expertise with regards to evaluation planning and methodology as well as expertise in the field of capacity development in developing countries.

The Evaluation Team is required to have:

- Proven capacity and extensive experience in management and conduct of complex evaluations,
- Strong understanding and experience in working with topics relevant to addressing capacity development issues,
- Strong understanding and experience in work involving partnerships and relationships with multilateral agencies, national/government agencies, civil society organisations and development organisations,
- Capacity and experience with evaluating policy dialogue and advocacy at national and international levels.

The evaluation team is expected to consist of minimum team members: two team members with international experience, who will be involved in all aspects of the evaluation, and at least one team member with substantial experiences from working with capacity development in developing countries. The team leader and team members are expected to complement each other so that the specific profile of the proposed team leader should have implications for the profiles of team members (and vice-versa). All suggested profiles will be assessed with a view to the role, competences and tasks they are suggested to cover in the team.

The evaluation team must include at least one member with knowledge of relevant Danish strategies and aid modalities. At least one team member must be able to read and understand Danish. A gender balanced team is preferable.

The tenders should clearly state who of the proposed team members cover which qualification criteria. The team must contain experiences with evaluation methodologies and tools that are proposed for the evaluation.

Qualifications International Evaluation Expert (Team Leader)

The Team Leader will be responsible for the overall management of the assignment, the team's reporting to and communication with Danida and will participate in meetings with Danida as well as with the Reference Group and other relevant stakeholder forums, as requested by Danida. The Team Leader is also responsible for the delivery of the outputs and thus should have experience in managing multi-disciplinary teams, producing high quality reports and working to meet demanding deadlines.

General qualifications

- Relevant higher academic degree. A profile with major emphasis on development issues, preferably with a minimum of 15 years of relevant and recent professional experience,
- Experience as team leader for multi-disciplinary teams,
- Adequacy for the Assignment,
- Extensive experience leading evaluations of development assistance with an extensive knowledge on and experience from evaluation approaches and application of evaluation methods, including theory based evaluations and contribution analysis (at least three substantial references as team leader for complex evaluations),
- Extensive knowledge of development policies, delivery mechanisms including support to multilateral agencies, civil society, harmonization and alignment, gender equality and types of modalities for development cooperation, including bilateral interventions,
- Proficiency in English,
- Experience from developing countries, including Danida partner countries.

Qualifications of the International Capacity Development Expert

General qualifications

- Relevant higher academic degree,
- At least 10 years of relevant professional experience from development cooperation; recent experience preferred.
- Adequacy for the Assignment,

- Substantial experience within design of policies and strategies for capacity development and implementation (at least four substantial references), preferably covering both bilateral and multilateral interventions.
- Substantial experience within application of methods for evaluation of capacity development, especially on institutional level, including use of drivers of change analysis, stake-holder analysis etc. (at least four substantial references)
- Experience with complex evaluations or larger reviews of development cooperation (at least two references).
- Proficiency in English.
- Experience from developing countries, including Danida partner countries.

Qualifications of Country Level Capacity Development Expert

General qualifications

- Relevant higher academic degree,
- At least 8 years of relevant professional experience from development cooperation; recent experience preferred,
- Adequacy for the Assignment,
- Specific profile on development and execution of policies and strategies for capacity development in developing countries, especially within public institutions, (at least four references),
- Other relevant experience with capacity development from country level interventions (at least two references).
- Experience with independent evaluations of development cooperation or larger thematic reviews (at least two references).
- Experience in the region and language
- Proficiency in English.

11. Requirements for the Consultant's Home Office Management, technical back-up, quality assurance (QA) and Business Integrity Management

The Consultant's home office support shall provide the following, to be covered by the Consultant's overheads:

- General home office administration and professional back-up;
- Implementation of the business integrity management plan as described in the Consultant's application for qualification and specified in the Consultant's technical tender;
- Quality assurance (QA) of the consultancy services in accordance with the Consultant's quality management and quality assurance system, as described in the

Consultant's application for qualification and specified in the Consultant's technical tender.

The technical tender shall include a detailed description of the proposed QA, to demonstrate the capacity of the Consultant to implement and verify a full QA programme. The Tenderer should select a person who is not a member of the evaluation team to be responsible for QA. The CV of this person shall be included in the technical tender. All QA activities should be well documented and be provided as part of abovementioned reporting of the evaluation process.

12. Eligibility

The DAC evaluation principles of independence of the Evaluation Team will be applied. In situations where conflict of interest occurs, candidates may be excluded from participation, if their participation may question the independence and impartiality of the evaluation. It is the responsibility of the bidders to inform the tender committee about any potential issues of conflict of interest. The final decision on eligibility, however, rests with the tender committee. Any firm or expert participating in the preparation or implementation of a project or programme directly related to the Denmark's capacity development funding may be excluded from participation in the tender, unless the involvement does not constitute unfair competition.

13. Annexes

Annex I: Approach Paper

Annex II: Specification of methodology

Danida Evaluation Department

June 2014

Annex 1: Approach Paper

Developed by Nils Boesen for the Joint Scandinavian Evaluation of Support to Capacity Development

1. Purpose and scope of the approach paper

This paper outlines key parameters of the joint approach to the evaluations of support to capacity development organized as parallel evaluations commissioned by Danida, Norad and Sida. The purpose of the paper is to guide the evaluation teams in their preparation of the on the overall approach; and to enable a constructive dialogue between the evaluators and the contracting agencies during the evaluation based on a shared general framework.

While the approach paper should guide the evaluations, it is expected and welcomed that evaluation teams suggest modifications and additions, which would add to the insights and robustness of the evaluations.

The Terms of Reference prepared by Danida, Norad and Sida, respectively, define the evaluation object, the scope and delimitations, specific deliverables and timelines, process and contractual aspects. The ToR take precedence over this paper.

After a short summary overview the approach paper outlines:

- A brief overview of the development of thinking of and approaches to capacity development (CD) among donors over the last decades;
- A basic analytical model and conceptualization of capacity, capacity development (CD) and support to CD;
- The key elements of a theory of change about CD that the evaluations will test;
- Elaboration of the four main focus areas of the evaluation(s);
- A brief conclusion

The approach paper is based on 'mainstream' literature on aid to support capacity development, guidance material, recent evaluations and meta-syntheses, including the literature review (Christoplos, Hedquist et al. 2014) made for this evaluation. Particular reference is made to the series of perspective notes published by OECD in 2011 ahead of the Busan meeting.¹ ().

2. Summary overview

The Scandinavian Evaluation of Capacity Development consists of three separate, parallel evaluations commissioned by Danida, Norad and SIDA, respectively. It is the desire to be able to draw conclusions across the evaluations of each agency, while each agency will also prioritise particular high-interest areas. A shared overall objective is to enable findings across the three evaluations, hence the joint overall approach outlined in this paper.

The evaluation will focus on interventions that have an *explicit* intention to support CD. It does not matter whether this intention is pursued with targeted inputs provided by the agencies (such as technical advisers, twinning, training etc.) or whether it is expected to happen by efforts of the partners

¹ <http://www.oecd.org/dac/governance-development/capacitydevelopmentourkeypublicationsanddocuments.htm>

or as an effect of the way support is given (e.g. (sector) budget support justified with its potential to support capacity development).

The evaluation will focus on a variety of modalities deemed significant by the respective agencies, from country/sector based support to regional/global programmes; as well as on a variety of instruments, e.g. twinning/peer based approaches, training, technical assistance and combinations of these and other possible means.

The focus of the evaluation is on capacity of the public sector, but that would include any interventions addressing factors and institutions outside the relevant institutions, if relevant for capacity development in public sector.

As specified in the Terms of Reference (ToR) the evaluation has both an accountability and learning focus. Within the overall key attention to *relevance* and *effectiveness* of the CD support from the agencies the evaluation will consider four issues:

- i. The relevance and opportunity of a “best fit” approach for CD support well adapted to specific intra- and inter-institutional dynamics and the wider context.
- ii. Within the “best fit” dimension, the appropriateness and legitimacy of external (donor) involvement in different dimensions of capacity development, and whether some processes may be so complex and demanding that the ability of donors to add value is limited.
- iii. The merits of looking beyond the supply side of public sector institutions to foster broader accountability relations or other types of collaboration with e.g. civil society, private sector, media or oversight institutions.
- iv. How a results-focused approach to aid for capacity development can serve to improve learning and accountability among aid agencies in future.

In the language of theories of change, the evaluation would thus start from the hypothesis that CD support from donors is (more) effective when it i) fits the drivers for and constraints to change (“builds on what is there”), ii) donor support is accepted as legitimate and appropriate; iii) uses results sensibly to measure progress, correct course and learn; and iv) looks beyond “supply-side” or “push” approaches that only work from the inside in public organisations. The evaluation will investigate if and how CD interventions adhere to this theory, and if and how the outcomes and impact of the CD and CD support confirms the theory of change.

3. Brief overview of past and current trends in CD thinking

From the 50'ies, development assistance had the ambition of delivering sustainable results (socially, institutionally, economically, technically, politically), and, by implication, foster endogenous capacities. The basic assumption was that capital investments plus transfer of knowledge/skills would suffice. Investment projects and technical assistance (TA) quickly became the dominant cooperation mode. The focus was squarely on the supply side provided by donor agencies – training and experts were the typical “soft” components added to the hardware. Implementation was in donor hands, assuming that learning/skills acquisition/institutional development would happen more or less by itself as a result.

Early reviews did not support that hypothesis – already in 68, the Pearson Commission found that “technical assistance often develops a life of its own, little related in either donor or recipient countries, to national or global development objectives”(Hradsky 2011), and later evaluations (Forss, Carlsen et al. 1988, Berg 1993, Land 2007, JICA 2008) have not found solid evidence of the broader effect of TA on sustainable capacity development. Attention also began to focus on the potential negative effects of TA (cost, distortions in local labour markets, disruptions in formal hierarchies, weak and twisted accountability mechanisms, distorted incentives through e.g. all kinds of salary supplements, and the proliferation of high-cost lifestyles among expats creating tensions with locals).

Despite the many negative reviews and evaluations of TA, the practice continued (and continues), often for reasons unrelated to capacity development purposes: TA served as implementation agents and safeguards for donor concerns – put bluntly, they were good at spending money and keep some degree of control over how funds were spent (Boesen 2001). Assuming double roles – both acting as implementation agent and as capacity development support adviser – often proved difficult.

Despite the introduction of the Logical Framework Approach in the mid-80'ies, with its insistence of starting from objectives and working backwards through outputs to activities and inputs, the focus for “institutional development” continued to be on the input or supply side. Demand was largely reduced to assumptions about political will and absorptive capacity which were rarely properly analysed, or for which systematic evidence were not sought.

In the mid-nineties, the disappointments with project-focused development assistance (“islands of success in seas of failure”) and TA led to two, in principle complementary, changes in approach among key multilateral and bilateral donors:

Firstly, programme-based approaches became in vogue (culminating with the Comprehensive Development Frameworks, and later the PRSPs) (Harrold 1995). Simply put, it was assumed that development efforts had to be nationally led and comprehensive, within somewhat consistent (sector) policy frameworks, and with a focus on a broad set of capacities, not least focusing on public financial management capacities, but also on service delivery and regulatory capacities. Results based management approaches came to the fore, as a belated reflection of the wave of New Public Management that had gained prominence in OECD countries.

Capacity was assumed to grow out of such programme approaches, which put the partners in “the driver’s seat”, as it was often phrased.

Secondly, as a complementary trend, capacity development became increasingly seen as a methodological discipline on its own, leaning heavily on mainstream approaches from organizational development, human resource development and management disciplines, increasingly tilting towards change management informed approaches. Some of the literature focused on what donors could and should do, but increasingly the mainstream view has been that CD is first and foremost an endogenous process, where outsiders can at best contribute – but never claim attribution (DAC 2006).

However, in practice – when programme based approaches was at its height and the pressures for parallel comprehensive reform efforts that would justify e.g. budget support were most intensive – there were rarely if ever room for the kind of CD support that the methodologies prescribed (e.g. ownership, incrementalism, flexibility, and attention to context, incentives, power, politics and

interests). Technocratic, linear blueprint approaches continued to dominate in practice, exacerbated by sometimes very rigid focus on results-matrixes and indicators.

Despite the increasing awareness – also translated into methodologies and guidelines (Boesen 2005, DFID 2007, Boesen 2011, Danida 2011) – that drivers and constraints to change, politics, power and interests matter for development and CD, the practice has seemingly only changed slowly and to a limited degree. The focus on the demand side is well established as theory, but it is difficult to discern it as a constituting element in the actual practice of development agencies.

Summing up, for the purposes of the evaluation approach: there has been a two pronged development: one displaying an ever more sophisticated understanding of change/development and the options and limitations of what outsiders can do to support change/reform. Buzz-words here for what outsiders can do are “facilitation”, “brokering”, “leveraging”, “complexity”, “flexibility” and, to some degree “emergence”(Morgan 2004, Andrews, Pritchett et al. 2012). And, on the other hand, an expectation that the broader approach to development assistance – programmes instead of projects, budget support instead of earmarked support – would by itself create space for endogenous CD processes – not so much by what donors would do, but maybe more what they would not do (e.g. distorting incentives in institutions, field TA, micromanage their assistance).

The translation of these two trends to actual practice – expressed in the Paris- and Accra-declarations - has at best been uneven, and the trends, in particular regarding programmatic approaches, may well have reversed. Fragile states may be the exception where at least in theory, harmonization and alignment is still in vogue.

Unaligned to the changes in declared approaches, methodologies and guidance described above, donors have on the other hand seemingly continued to pay special attention to those aspects of capacity that were deemed to be of particular importance to the effectiveness of their aid. Policy making, financial management, results based management, monitoring and evaluation as well as anti-corruption and transparency capabilities have ranked high (Independent Evaluation Office and Operations Evaluation Department 2004, Operations Evaluation Department 2004, Boesen and Dietvorst 2007), while broader downstream service delivery capacities (e.g. development effectiveness capacities which are more comprehensive than aid effectiveness capacities) may not always have received the same attention.

4. Concepts and definitions

The CD field uses broadly defined concepts, and there has never been a shared, precisely defined vocabulary. This is, however, not very different from the situation in the academic fields of organisations, institutions and systems where a variety of concepts are used without sharp delimitation of meaning.²

² There has been numerous calls for getting to sharper, and shared, concepts. That has for obvious reasons not been successful, but much time has been spent on battles over more or less subtle definitional tweaks. E.g. the difference between *capacity* and *capability*, between *latent* and *actual* capacity, between *organisations* and *institutions*, and, maybe most frequently, between *capacity development* and *capacity building*. Opponents of using the latter – often Europeans – argue that a building metaphor is irrelevant because capacity develops much more organically, and not by engineered design. North Americans often prefer capacity building .

OECD/DAC's definition from 2006 (DAC 2006) is the accepted common definition: *Capacity is understood as the ability of people, organizations and society as a whole to manage their affairs successfully. ... 'Capacity development' is understood as the process whereby people, organizations and society as a whole unleash, strengthen, create, adapt and maintain capacity over time.*"

Evidently, defining "capacity" as "ability" is on the edge of being tautological. More important is the mentioning of "people, organisations and society as a whole", often translated into a request that CD must focus on individuals, organisations and (broader societal) systems or institutions to be effective.

On "capacity development", the important point of the definition is the focus on CD as a *process*, and the focus on people, organisations and societies as at same time actors and beneficiaries of CD. That is, the definition of CD alludes to CD as (largely) an endogenous process – like learning, it can be supported (e.g. by teaching), but the teacher cannot "learn" the student anything – the learning (or change) takes place inside the individual (or organization, or society), and so does CD eventually, because the resulting capacity resides in people, organisations and societies.

Attempts to provide alternative definitions often run into problems because they include normative or prescriptive elements or a whole theory of change about how CD happens.

This evaluation will therefore stick to the reasonable and intuitively shared idea about "capacity" as the ability to perform ("perform" and "performance" are other terms which have no agreed precise definition). A basic understanding of CD as a process eventually resulting in changed capacity in individuals, organisations or systems/societies is on the other hand critically important because it moves the focus from the teacher (or CD supporter) to the learner (the person or organization or system developing capacities).

Moving beyond the definitions, the next crucial matter is to identify what the desirable *elements or factors* of capacity could be (for different persons, organisations, systems, and for different contexts). A presidential office presumably needs a different set of capacities to perform well than do a primary school, though factors or elements as management, structure, systems etc. are present in both type of organisations. In addition to the capacity factors, a theory of change about capacity development have to identify the *attributes of capacity development processes* are when these are successful – or, reversely, when they are not. That is, a top-down and unconsultative approach to change may work better in the presidential office than in a school. These are the fields where hypotheses and evidence speak, rather than in battles of terminology.

This paper will return to these to critical issues after outlining a generic analytical model for CD.

5. The analytical model – the generic cause/effect chain

This section presents an analytical model of organisations and change that is descriptive and thus does not provide answers to which capacity factors to strive for or how to design change processes – but which allows a structured discussion about this in a simple cause-effect chain. The analytical model builds on one of the most frequently (but not always consistently) applied perspective on capacity development and organizational development: *an open systems approach* which sees organisations (and people, and systems) as open in the sense that they have permeable boundaries and constantly interact with their environment, influencing it and being influenced by it (Harrison 1994, Harrison and Shirom

1999, Boesen and Therkildsen 2004). The generic logic is as follows, at this stage for how capacity leads to impact, without considering the context or environment:

3. Capacity -> 4. Performance/Outputs -> 5. Outcomes -> 6. Impact

Note that this is not a project logic, but a generic logic for how people, organisations or systems use their capacity (e.g. a rural health clinic) to perform/deliver services (attend pregnant women) with an outcome (more women give birth in clinics) with an impact (lower maternal death incidence). In this logic, the outputs of capacity is a (continuous, adaptable) *supply* of services (or products, or enforcement of regulations). The outcome is the *use* of these services/products/compliance with regulations (or the *demand* side), leading to an ulterior impact.

Adding the environment or context, the above logical chain underlines that capacity is used (and changed) in a context which it is influenced by and which it – to a certain degree - can influence. Staff paid according to government rules may for example have such a meagre salary that they only pretend to work (“the state pretend to pay us, we pretend to work”) – or really only perform when they have other incentives to do so. The same relation to the environment goes for outputs, outcomes and impact – they are delivered and used in a context, and both outcome and impact depend on context factors (e.g. if women have no means of getting to the health clinic they cannot use the service).

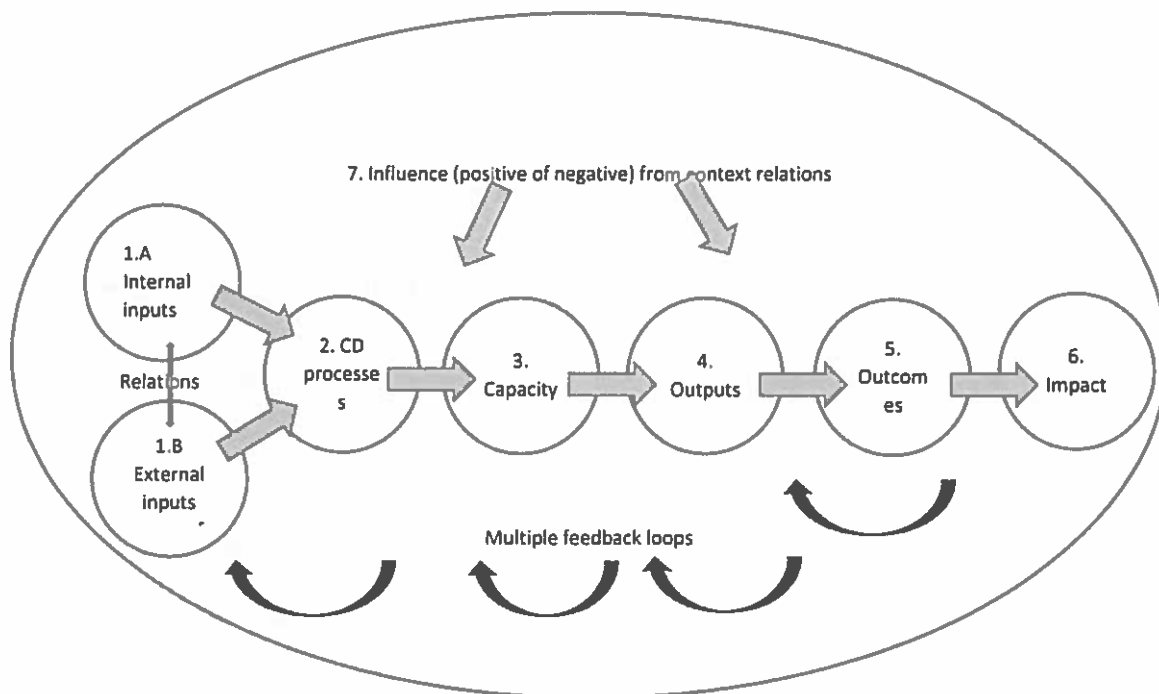
To add the CD perspective, the logical chain expands as follows:

1. Inputs to CD -> 2. CD processes -> 3. Enhanced Capacity -> -> 4. Enhanced Performance/Outputs -> 5. Enhanced Outcomes -> 6. Enhanced Impact, all in 7. Context within and beyond influence.

The figure on the next page illustrates the analytical model, distinguishing between inputs to CD from internal actors and external actors, respectively. This allows a separate look at donor support to CD, and at the relations between the donor(s) and those whose capacities are to be developed.

Note that this logical chain is purely conceptual, and therefore referred to as an analytical model rather than a Theory of Change. It does not yet specify a thesis about which inputs to CD (internal or external) that will work, what good or less good CD processes are. It does not detail the factors/element of capacity that are important: it may be “functional capacities” and systems (e.g. a new IT system and the skills to operate it), or more value-based elements (management style, transparency, participation) – and it may well be contested which elements that are important or most important. Better performance is not prescribed either, it can be quantitative and/or qualitative enhanced services/products/regulations. The analytical model does not specify which context factors that are important for success or failure, and if or how they may be influenced by change agents or other stakeholders. It simply offers a frame for this discussion, arguing that the logic behind CD and CD support goes from inputs to CD processes to ulterior impact.

Generic Search Model for Cause-Effect Relations in Capacity Development



The analytical model thus focuses on enhanced capacity, which is the intended result of deliberate CD processes. But even that is not at all a given – capacity enhancement may come around “by itself”, without any deliberate or identifiable processes or actors, or with only very limited interventions that claim to be driven by a CD ambition. Using biological metaphors, a plant may grow simply because the environment provides rain and fertile soil, it does not need an expert to tell it how to develop. In the CD literature, this possibility has often been described as “emergence”. Some have even claimed that it is the dominant way in which capacity develops. This is in clear opposition to those with a more engineering approach who claim that a clear, meticulously planned, measurable, results- and analytically based approach can develop capacities. In this evaluation, both elements of emergent and planned capacity enhancement should be identified – but only in the context of CD support that had the intention to develop capacity.

To make the model appropriate for the special case of development cooperation, it is furthermore useful to distinguish between organisation-internal investments in CD, and external support, respectively. This allows looking separately at the contributions of different actors, as well as the relations between them, thereby allowing disentangling the issue of ownership of change processes. In specific terms, this would raise questions about who are doing what, who are deciding what, and in terms of the quality of relations: are the parties seeing each other assuming appropriate and legitimate roles, are they effectively agreeing on objectives of the CD processes, and how should this process be conducted; etc.

The resulting picture, as illustrated above, is the proposed analytical model for the Theory/ies of Change that the evaluation will research.

As underlined, the model is (largely) “empty” except for some fundamental assumptions about organisations (open systems). What the requisite capacity for a given organization or system would consist of, is not detailed.

The model on the next page is largely identical with the model behind the Results-Oriented Approach to Capacity Change (ROACH) approach developed by Danida in 2003-2005 (Boesen and Therkildsen 2004). It allows investigating changes in each of the circles, as well as on processes and relations. This is not only relations between possible internal and external inputs to CD processes, but also the relations between the organisation (or system of organisations) and the context, from the formal governance structure to relations to suppliers, to users/consumers and to other relevant stakeholders. Further, as recommended in the ROACH, it allows the description of changes over time – and then afterwards to consider whether the changes in one element can be ascribed causally (partially) to change in another element, both in the linear flow and in relation to other explanatory factors in the context (Boesen, Christensen et al. 2002). Indicators and means of verification of inputs, changes in capacity, outputs, outcomes etc. are technically readily available, though in practice hard to collect ex post as they are rarely collected at the beginning of capacity development support. However, any evaluation of CD faces a paucity of available data, no matter which model is applied, as will be further discussed later in this paper.

6. A Theory of Change for Capacity Development

What are, then, the *elements of capacity* that are most relevant to a given organization in a given context? What are the *characteristics of the CD processes* actually supported by or imposed by donor agencies? What were they supposed to be (according to methodologies, or an idea of good practice)? What is the balance between internal and external investments in change, what are the relations between those involved (ownership, leadership, resource control etc.)?

In the capacity development literature referred to above there is not consensus on which capacity elements that are most crucial for the performance of an organisation. As briefly discussed above, some donors may tend to focus excessively on the formal capacities that ensure that aid is efficiently used (funds spent on time for the agreed budget items) and has an immediate effect (the envisaged direct outputs have been produced – a bridge constructed, x number of teachers can teach a new math curriculum). This is a narrower set of capacities than what is required to achieve effective development or impact (e.g. ensuring that pupils learn math relevantly when they are in school requires many other things than good financial management and teachers being able to teach the curriculum: parent support, textbooks, incentives for teachers to show up, reasonable student/teacher ratios etc.).

In practice, donors may also consider that they have little to offer regarding the softer, informal and more 'political' capacities that their partners need to perform in the context, such as good management, capacity to reach out to stakeholders, and capacity to stay on good terms with the most powerful among the stakeholders. The dialogue between donors and partners may skip such items, which can be sensitive, leading to agreements about the need for "functional" capacities only, even if both parties are aware that they are not sufficient.

There is broader consensus about the parameters that influence effectiveness and efficiency of deliberate CD *processes*. That is, there is implicitly, a Theory of Change about Capacity Development as it can be read out of recent guidance from the European scene (Boesen 2005) and the UN system (UNDP 2007), as well as evaluations. This Theory of Change can be summarized as follows:

- Capacity development is largely an endogenous process, strongly conditioned by the structural, social and institutional drivers of and constraints to change which changes over time. CD and external CD support is more likely to be successful when it builds on what is there, and is driven, managed and owned by local stakeholders in and around the relevant organisations.
- Successful CD processes tend to be adaptive and flexible, even in terms of moving goalposts, as CD nearly invariably takes place in a context of messiness and complexity, where a mix of formal and informal governance and accountability mechanisms (including loyalty based and patrimonial mechanisms) shape incentives. Successful CD processes tend to work on a mix of both internal and external factors, and on both formal and informal, and functional and political aspects of the organization(s) and their immediate context.
- Results can meaningfully be measured both at the level of changes in capacity (better systems, more efficient structures, better performing individuals, better communications, better adaptation and resilience etc.) and at the level of organisational performance (changes in outputs, whether quality, quantity, cost or relevance or a combination of these).

- External support can, consistently with the above, expectedly only *contribute* to capacity and performance enhancements because so many other factors are in play to make CD success. They could still make the critical difference, but other forces have to push in the same direction. Relational issues between development agencies and partners in relation to CD support is therefore of critical importance, including the perceived legitimacy of each partner's actions or inactions in the eyes of the other partner.
- A clear focus on achieving measurable changes in organizational performance is useful. This aspect, which is one of the four focus areas of the evaluation, is further discussed in section 8 below.
- External inputs like training, TA and "knowledge transfer" need intrinsic qualities to be effective, but their effectiveness depends most of all of the strength of demand from the partners³.

This Theory of Change for Capacity Development is a hypothesis only. The evaluation will focus on specific sub-sets of questions to test whether practice aligns to this theory, and whether or not there is a correlation between interventions that align to this theory and the effectiveness of the capacity development support provided by donors.

7. Unfolding the "best fit" and legitimacy focus areas: complexity of capacity, adequacy of CD processes and donor responsiveness

What does it mean to hypothesize that a "best fit" support from donors to CD is likely to be more effective than a "blueprint approach" where a purportedly "best practice" is sought introduced as a blanket replacement of the existing capacities? While the idea may sound intuitively right, it is helpful to break it down into key dimensions that can be identified and measured during the evaluation. This also helps avoid stereotyped discussion - there may be many ways to skin a cat, but applying some international standards, e.g. measuring if a science laboratory is performing acceptably, is done in one, standardized way, according to what can rightly be called a blueprint – or best practice standard. On the other hand, there is no blueprint for a "best" political system – nor for how a political system is developed or changed. The same go for CD processes – there is no "best practice" for how to conduct reform processes or major organizational change processes.

A key dimension of relevance for the "best fit" in relation to capacity is the level of *difficulty or complexity* of the capacity itself. Thus, the capacities required to produce simple iron rods are less sophisticated and simpler than capacities required to produce computer processors.

A number of factors related to the "nature" of the capacity itself influence the difficulty or complexity:

- the *specificity* of the products/services that the capacity will be used to produce/deliver (Israel 1987). Higher specificity eases CD (e.g. the capacity to deliver a school lunch is easier achieved than the capacity to produce internationally approved PhD graduates). "Capacity for good

³ Peer learning approaches, and some 'South-South' approaches, are arguably more effective because they fit demand better and are thus in this way of higher relevance and quality – but the presence of demand is the key issue for making any supply effective.

“policymaking” is on the other hand not a set of easily specified skills and business processes, while e.g. simpler surgery is performed according to very specific, standardised processes.

- The *scope or distribution* of the capacities. A complex task like getting monetary policies right can be successfully performed by a small group of highly qualified economists, provided they have the support of the government leadership. Such capacity can rather easily be developed (or outright bought) – there are several examples of countries “fixing” their monetary (and also fiscal) policy by recruiting small teams of economist from the diaspora. On the other hand, getting quality education in most classrooms in a country – a very complex and widely distributed task - is an enormous capacity development challenge, depending on various systems governed by widely different governance mechanisms (from teachers to textbooks and buildings, to budget allocations and parent involvement – etc.). Conversely, producing school meals at all schools or conducting a vaccination campaign demands much “simpler” capacities and CD processes, even if it is across a country, because it does not depend much on other organisations or systems.
- The *incentives to performance*. At the level of staff this is often about a combination of remuneration, status and recognition, perceived risks and gains in relation to these factors, and intrinsic motivation. Typically, in many developing countries, the combined incentives can be weak at a systemic level (salary conditions in the public sector, perceptions of risks and opportunities) – making CD efforts targeted at individual organisations susceptible to failure because the general context conditions for successful CD are not conducive.
- The *interests of elites*. New capacities to e.g. tax the wealthy, or make land markets transparent and competitive, threatens interests of powerful elites. The capacities in question may be simple or complex, but the fact that they may not be in the interest of elites will make them much harder to develop. Both CD and reforms are generating winners and losers. The more CD and reform threatens the power and interests of elites (inside and around an organization, in a community, in a sector or at national level), the more resistance will change be met with and the more difficult will it be to develop capacities and transform it into performance (Robinson 2006).
- The *affinity* of the capacities (and products/services) with the dominant social and cultural norms and values. Going against the deeper grain of a society is simply more difficult than going with it. Examples abound around e.g. gender equality objectives which, despite declared intentions, often move much slower ahead than wished for.

In summary: Capacity development is easier the more specific the capacities, the less the systemic dependencies, the stronger the incentives, the less against powerful elite interests and the more in tune with the grain of society. A prime example: the capacity to use cell phones far beyond their initial focus has grown largely by itself, quickly and easily. On the other hand, despite massive donor support over 25-30 years, public financial management in many African countries still displays serious weaknesses.

The complexity and difficulty of developing certain capacities has often led donors to pursue “paths of least resistance”. They seldom if ever meet opposition to support that will increase budgets and staff, or provide new buildings, infrastructure, computers and vehicles. Training and various forms of study and knowledge acquisition options are also most often welcomed, and even sought for. The problem with these forms of “simple” support – which have been the backbone of donor approaches to CD since the

60ies – is that they have generally not proven effective, because there were also other capacities needed – those harder to develop.

Thus, when it comes to changing priorities, incentives, management or organizational culture this will most often meet stronger resistance, posing much bigger demands to the capacity development or change *processes*. To put it simple: if a certain new stage of capacity is complex and/or difficult to reach, it demands more power (effort, time, resources, allies...) of the capacity development process to get there.

The change (and reform) *processes* (with or without the participation of a donor) can be assessed on their *adequacy* to the task in hand – that is, the adequacy to the complexity of the capacities to be developed:

- *Level of effort*: Depending on the complexity of the capacity to be developed, and the process required to achieve it, the combined efforts of the partners and the donors may be grossly underestimated, relatively balanced or apparently too much and/or too costly. Most often, evaluations report that change and reform processes underestimated the time and effort required to push through resistance, upgrade capacities and transform it into performance.
- *Availability of champions*: Champions are often identified as external leaders with the power to get support for the change and overcome (or, as things go, adapt to...) the resistance. While there is often a focus on individuals, it may be wiser to look for stakeholder coalitions and groups of champions whose combined power and engagement is bigger than the power of those resisting change.
- *Change agents*: Credible change processes need credible and legitimate change agents, with access to bosses, networks, technical, managerial, political, communicative and financial resources. Donors often come in here, supplying technical and financial resources – but it may be in a poor match with the other resources needed, and e.g. donor-contracted technical assistants acting as de facto heads of programme management units often lack the legitimacy and ability to act on the informal lines.
- *Time-horizons and rhythms of change processes*: Time-horizons can be overtly short, or far too long – and the rhythms would usually give room for victories, highlights, pauses and adaptations to a changing context.
- *Big bang or incremental approaches*: Linked to the previous bullet, CD and reforms can adopt all-in-one-go approaches or sequence and scope reforms in stages. Most literature would say that successful reforms most often follow the latter approach.⁴
- *Clarity of vision and results*. There is little evidence in literature that CD and reform normally follow the proclaimed intentions, and little advice that this should be the recommended course of action. But there is some evidence that a level of obsession with progress and results by the key change agents and sponsors helps (Collins 2001). This should not be confused with formal results-matrixes.

⁴ The incremental approach was tested in a PFM reform in Bangladesh some years ago, where a number of successive “platforms” of reform achievements was defined ex ante, with the idea that the reformers and their donors would only proceed to the next platform once the previous one had been reached. In PFM support to at least Mozambique is was at some point around 2000 argued that reform should start with “simple” processes before addressing the more “difficult” ones.

Finally, the donor support should be assessed by its *responsiveness* to the situation, including the following:

- *The degree of legitimacy of the donor intervention, and 'donor-steering'*: This sensitive, but crucially important issue is often phrased in terms of the level of ownership by the partners, and often left aside without a more refined analysis or attention. A more detailed analysis would look at to what degree the donor has been invited as a legitimate partner, and to what degree the donor adapts to the rules of the house, and what the incentives for the partner are to invite the donor into the CD process.⁵ This is maybe one of the most important classifiers of donor support to CD, catching the degree to which donors steer. In the extreme, donors can attempt to singlehanded drive reform through project documents they have written themselves, a Project Implementation Unit, donor-recruited TA etc. At the other extreme they can offer passive accompaniment to a pot of money put at the disposal of the partner to achieve (or for achieving) pre-defined CD results specified by the partner. In between, but towards the less intrusive, catalytic and leveraging approaches (e.g. bringing stakeholders together, arranging (South-South) exchanges; making knowledge/experiences available, supporting local CD institutions) have gained prominence in recent years.
- *Grasping the setting*: Often linked to the bullet above, donors may put few or many resources into understanding the setting (all the parameters outlined above in relation to change and change processes). Getting to a "best fit" approach evidently requires a good understanding also of what is going on "behind the façade".⁶
- *Clarity of vision and results*: Earlier evaluations and portfolio screenings have quite consistently found that CD has been under-specified to a degree where it has been difficult to assign any real accountability for results afterwards.
- *Level of attention*: While the initial specification of CD is often low, the follow-up on CD support and results (evidenced in reports, changes to approaches/activity plans, budgets etc.) may often be conspicuously absent – for reasons that may reflect the initial lack of shared aspirations and clarity of roles, processes and main results.
- *Incentives*: Donor staff and representatives face multiple incentives in their relations to their own organisations and to the partners. Incentives can put disbursement higher than CD support; compliance higher than real results; risk aversion higher than risk taking; and short term ease of doing business over getting things right for the long term.

While there is no simple typology with distinct "species" of CD and CD support coming out of the lists of significant factors above, there are three important dimensions of CD processes and donor support, respectively, that will shape if and how the donor support is a good fit to the situation:

- the *complexity* of the capacity or performance strived for;
- the *adequacy of the CD process*; and
- the *responsiveness of donor support*.

⁵ An African minister was once quoted for saying that "capacity development was like having donors in the bedroom".

⁶ The Dutch Ministry of Foreign Affairs coined this term for their efforts to understand the political economy in their partner countries (Harth and Waltmans 2007).

The theory would be that the higher the complexity of the capacity to be developed, the more sophisticated should the approach to change be, and the higher the required responsiveness of the donor support.

These dimensions (and the underlying factors) thus serve as a framework for assessing the “best fit” question which is a core theme of the evaluation.

By looking at these dimensions the evaluation may cast light on whether some CD ambitions are so complex that the ability of donors to respond is simply not available. The tools at hand (externally acquired resources, dialogue, presence, money) and the position of donors (outsiders, operating in a mix of own domestic politics, diplomacy, politics in the partner country, and international development fashions and fads) may simply not be adequate to contribute effectively to complex, path dependent and messy development processes whose time perspective is several decades rather than short term donor cycles (Booth and Cammack 2013).

In the evaluation language, the above discussion touches in particular on *relevance* and *effectiveness* of Nordic CD support. Relevance is thus not about the need – any organization can in principle get better, anytime, everywhere – but how the donor support fitted to the situation and whether that lead to the intended results.

8. Unfolding “looking beyond internal capacity” and “working with results” focus areas

Looking beyond internal capacity

Traditionally, donor support to CD has targeted the “internal machinery” of public sector institutions. New skills, systems, organizational infrastructure, strategies etc. etc. were assumed to address the “capacity deficit” and thereby lead to enhanced performance. It was, implicitly or explicitly, assumed that the governance arrangements around public sector organisations that shaped incentives were either grossly adequate, or, if inadequate, then beyond the reach and influence of the organization itself and the donors supporting it. Attempts to address public sector wide incentive problems (through salary decompressions; merit-based recruitments and promotions; financial management improvements and anti-corruption initiatives etc.) were sought addressed through national public sector reforms.

This basic perception that “supply-side” CD support to the internal workings of individual organisations would be effective was increasingly questioned through the 90’ies. The 2004-World Development Report (World Bank 2004) summarized this new look introducing a strong focus on the accountabilities and governance mechanisms of public sector organisations, in particular those charged with delivering basic services. The main argument was that the formal, top-down “principal-agent” mechanisms where politicians and ministers were supposed to hold front-line service providers accountable were not effective. This “long route accountability” should be complemented by “short route accountability” where the services users would also act as principals holding providers to account. Water Users Committees, Teacher-Parent Associations, citizen scorecards and publishing of budgets, accounts and results were the key ingredients in a wave of attempts to “work from the outside-in” – or “pull-approaches”, as complementary alternatives to “push-approaches” working from the inside. These

approaches also found their way into CD guidance of e.g. Danida, the European Commission, ADB and others (Boesen and Therkildsen 2004, Boesen 2005, ADB 2006).

The present evaluation will look at if and how CD support from the Nordic countries has sought to support CD processes not only from the inside, but also through such “pull-approaches” that would strengthen oversight, accountability and transparency. This could be through capacity development support to external stakeholders that should and could oversee and hold public sector organisations to account, thereby providing incentives for these organisations to perform better. Apart from direct users of services it could be oversight institutions such as General Auditors, elected councils and anti-corruption bodies.

It should of course not be assumed *ex ante* that such support would have the desired effects – recent research (Booth and Cammack 2013) has been questioning the effectiveness of such measures, in particular if they are seen as an alternative, rather than a complement to other CD measures. The focus in this evaluation is thus not at all to see “pull-approaches” as a possible “magic bullet” for CD, but to ascertain if and how CD support has departed from a comprehensive analysis and understanding of the multiple internal and external factors that shape capacity and performance.

Working with results

The fourth focus area of the evaluation approach is how CD support from the Nordic donors has worked with results and indicators – how were these formulated up front during design or inception phases; were baselines identified; and how was the follow up in reporting and in terms of adjustments of plans and future results-planning?

This focus area touches a much larger and sometimes very heated debate about the merits of results-based management (RBM) as this came to the fore as an essential part of the New Public Management (NPM) wave that started in OECD countries in the 80’ies. RBM was seen as a means to displace old-fashioned rules-and-regulation based public administration with a system that would give managers concrete targets to achieve – and more freedom to choose how to achieve them. Academic reviews have clearly demonstrated that this is much easier said than done – RBM has advantages but also disadvantages (Hood 1991, Pollitt and Bouckaert 2004), and NPM-type reforms may be particularly difficult to pursue in developing countries (Schick 1998, Schick 2004).

The focus here is, within this broader RBM perspective, more narrowly on how CD support has actually used results and indicators – or not done so. All Nordic agencies have had and has a focus on results and evidence – but particularly in relation to CD it has been a frequent finding that results are not used consistently, and not in accordance with the stipulations of e.g. the standard Logical Framework Approach prescriptions which Norad spearheaded in the 80’ies and both Danida and Sida have subscribed to (NORAD 1990, Danida 1996, Boesen, Christensen et al. 2002, Norad 2008, Kruse and Forss 2014).

The frequently observed paucity of evidence when it comes to the effectiveness of CD support is part of the rationale for this focus area of the evaluation. Looking at how results have been planned, and how evidence has been collected is expected to cast further light on whether this paucity can be confirmed, and why results and evidence are used – or not used – the way they are in CD support.

Results orientation is indeed useful for CD. Specific capacity and performance results serves clarify directions for those involved. Explicit ideas of cause-effect relations – that is, results at different levels as per the analytical model in this paper or a similar model – would help focus minds and actions of the important things that will have the desired effects, rather than unimportant (but maybe uncontroversial) activities that do not add value to the capacity or performance of an organization. Monitoring the progress would be essential for meaningful learning to take place, enabling managers to modify processes and align expectations to current realities.

It has been pointed out that results-orientation as a formal system only may not have effects except reinforcing a “tick-the-box” culture. If a donor intervention is not perceived as legitimate by a partner, and if the partner and/or the donor does not have an organizational culture and a management that is results-focused, then formal requirements (often posed by the donors) should not be expected to work. Literature on the other hand indicates that impatient leaders with a personal drive for achieving daily results are better at transforming organisations (Collins 2001). The evaluation should look at results-orientation around CD support also in this broader perspective, looking at the availability or lack of formal instruments of RBM and evidence in the wider context of whether and how results matters to leaders and organisations.

9. Conclusion

The approach paper has identified broad definitions of CD taken from OECD/DAC and widely accepted among DAC members. It has outlined an analytical model that requires the evaluation teams to distinguish between inputs to CD, CD or change processes, the intended resulting capacity, and the performance and wider outcomes that this capacity would contribute to. Within this analytical model the approach paper has hypothesized a Theory of Change that reflects mainstream views about capacity development: that it is more relevant and effective when it seeks an optimal fit to the context; when the role donors play is legitimate; when the CD process looks beyond the internal factors in public sector organisations; and when there is a sensible results-orientation.

The core evaluation questions outlined in the TOR also points beyond the four focus areas outlined in this paper, and it is expected and welcomed that the evaluation will identify other key attributes of successful CD support or key attributes of unsuccessful support. The approach paper sets a basic analytical frame for looking at and a analyzing the evidence, and for addressing some of the key tenets of contemporary thinking about capacity development.

Bibliography

- ADB (2006). Integrating Capacity Development Into Country Programs and Operations: Proposed Medium-Term Framework and Action Plan 2006–2010. Manila, ADB.
- Andrews, M., L. Pritchett and M. Woolcock (2012). Escaping Capability Traps through Problem Driven Iterative Adaptation (PDIA). CID Working Paper No. 240. Boston, Harvard University.
- Berg, E. (1993). Re-thinking Technical Cooperation. New York, UNDP.
- Boesen, N. (2001). Advisers in Danish Bilateral Development Co-operation. Copenhagen, Danida.
- Boesen, N. (2005). Institutional Assessment and Capacity Development: Why, What and How? Aid Delivery Methods - Concept Paper: a Contribution to Development Thinking. Brussels, EuropeAid.
- Boesen, N. (2011). Perspectives Note: The Enabling Environment for Capacity Development. Perspectives Notes on Capacity Development. J. Hradsky, N. Boesen, A. Land et al. Paris, OECD.
- Boesen, N., P. F. Christensen and O. Therkildsen (2002). Capacity Development Evaluation. Copenhagen, Danida.
- Boesen, N. and D. Dietvorst (2007). Sector Wide Approaches in Motion: From an Aid Delivery to a Sector Development Perspective. Roskilde, Nils Boesen Process & Change Management.
- Boesen, N. and O. Therkildsen (2004). Between Naivety and Cynicism: A Pragmatic Approach to Donor Support for Public-Sector Capacity Development. Copenhagen, Danida.
- Booth, D. and D. Cammack (2013). Governance for Development in Africa: Solving Collective Action Problems. London, Zed Press.
- Christoplos, J., A. L. Hedquist and K. Engstrand (2014). Literature Review for the Joint Evaluation on Capacity Development (Danida/Norad/Sida), Indevlop.
- Collins, J. (2001). Good to Great. London Random House Business Books.
- DAC (2006). The Challenge of Capacity Development: Working Towards Good Practice. Paris, OECD.
- Danida (1996). LFA: a Flexible Tool for Participatory Development. Copenhagen, Danida.
- Danida (2011). Addressing Capacity Development in Danish Development Cooperation.
- DFID (2007). Drivers of Change: The Politics of Development - Briefing Note. London, DFID.
- Forss, K., J. Carlsen, E. Frøyland, T. Sitari and K. Vilby (1988). Evaluation of the Effectiveness of Technical Assistance Personnel, Danida, Sida, Norad, Finnida.
- Harrison, M. I. (1994). Diagnosing Organizations: Methods, Models, and Processes. Thousand Oaks, SAGE Publications.
- Harrison, M. I. and A. Shirom (1999). Organizational Diagnosis and Assessment: Bridging Theory and Practice. Thousand Oaks, SAGE Publications.
- Harrold, P. (1995). The Broad Sector Approach to Investment Lending: Sector Investment Programs. World Bank Discussion Papers. Washington D.C., World Bank.
- Harth, Karen and Jan Waltmans. 2007. "Behind the Façade - the Informal Reality in Developing Countries. DEK, the Hague.

- Hood, C. (1991). "A Public Management for all Seasons?" Public Administration 69.
- Hradsky, J. (2011). Perspectives note: Technical Co-operation for Capacity Development. Perspective Notes on Capacity Development. J. Hradsky, N. Boesen, A. Land et al. Paris, OECD.
- Independent Evaluation Office and Operations Evaluation Department (2004). Republic of Mozambique: Evaluation of the Poverty Reduction Strategy Paper (PRSP) Process and Arrangements Under the Poverty Reduction and Growth Facility (PRGF). Washington D.C., World Bank and IMF.
- Israel, A. (1987). Institutional Development. Washington, World Bank.
- JICA (2008). Joint Study on Effective Technical Cooperation for Capacity Development: Synthesis Report.
- Kruse, S.-E. and K. Forss (2014). Methodological approaches to evaluate support to capacity development, Nordic Consulting Group.
- Land, T. (2007). Joint Evaluation Study of Provision of Technical Assistance Personnel. Maastricht, ECDPM.
- Morgan, P. (2004). What is Capacity? Going beyond the Conventional Wisdom. Maastricht, ECDPM.
- NORAD (1990). The Logical Framework Approach. Oslo.
- Norad (2008). Results Management in Norwegian Development Cooperation. Oslo, Norad.
- Operations Evaluation Department (2004). OED Review of the Poverty Reduction Strategy (PRS) Process. Washington D.C., World Bank.
- Pollitt, C. and G. Bouckaert (2004). Public Management Reform - A Comparative Analysis. Oxford, Oxford University Press.
- Robinson, M. (2006). "The Political Economy of Governance Reforms in Uganda." IDS Discussion Paper 386.
- Schick, A. (1998). "Why Most Developing Countries Should Not Try New Zealand Reforms." The World Bank Research Observer 13(1).
- Schick, A. (2004). Is Governance Reform in Advanced Countries Relevant to Developing Countries? Conference on the Political Economy and Governance Reform. Copenhagen.
- UNDP (2007). Supporting Capacity Development: The UNDP Approach. New York, UNDP.
- World Bank (2004). Making Services Work for Poor People, Oxford University Press.

Annex 2: Specification of methodology

1. Introduction

As specified in the TOR, the evaluation commissioned by each of the agencies Danida, Sida and Norad will comprise:

- A portfolio screening of a larger sample of capacity development (CD) interventions¹
- A desk-based review of a smaller sample of interventions
- Country studies

This annex specifies the purpose, scope, sampling criteria and methodology of the broader portfolio screening and the desk-based review, and outlines key methodology elements of the subsequent country studies.

2. Portfolio screening

The portfolio screening serves the following purposes for the evaluation:

- Inform the choice of cases for the desk-based review as well as country studies, and inform the refinement of the methodology approach for these steps.
- Provide findings to be reported and analysed in the final evaluation report as to what CD interventions consist of and other typical features of CD interventions indicated below.
- Allow comparison and subsequent statistical analysis across the evaluations by each Nordic country, possibly including the sharing of data with other researchers for further statistical analysis beyond these evaluations.

The portfolio screening starts with a database search, supplemented by inputs from the client, to identify a relatively large sample (30-50) of interventions with proclaimed CD objectives and/or components. It will be based on the following criteria:

- Capacity development as explicit intention, whether alone or as part of broader objectives.
- A description of intentions is available at a level of specification as in typical programme/project documents.
- Initiated at least three years ago, and if completed, then completed not more than three years ago. The selection should aim at a mixture of relatively new initiatives, and interventions that have been going on over some time.
- Geographic and thematic/sector spread as deemed relevant by each agency.

¹ The term 'intervention' is here used for any closely coordinated set of aid financed activities explicitly aiming to support capacity development, in line with the definition under 'scope' in the Terms of Reference. Thus, 'intervention' may sometimes equal projects or programmes, but since efforts to support capacity development may often be integrated in programmes aiming at other objectives, it may also be limited only to a certain set of activities within projects or programmes. It may also refer primarily to the choice of modality for how aid is delivered (e.g. budget support) within projects or programmes if justified as support capacity development.

- Interventions that are of a size/salience that make them indicative of the “CD portfolio” of the respective donor country.

The screening will be based on a limited number (1-3) of documents per selected intervention. This may be pre-implementation documents (identification/feasibility reports, programme/project documents, appraisals) and/or, if completed, end reports or similar. The combined level of effort would not exceed 2-4 hours per sampled case.

This phase will mostly involve collection of descriptive data. The data collected will be tabulated in electronic format (a template excel sheet or database will be made available). A scorecard (drawn directly from the tabulations) will be made on each intervention together with a summary tabulation of quantitative parameters across the screening (with spread and averages), and a max. 10 page narrative report. The report with summary tables will be annexed to the final evaluation report.

3. Desk-based review

The desk-based review will provide additional substantiation by validating, adjusting and deepening the results of the initial screening. In addition to the mostly descriptive information from the portfolio screening, the desk-based review will elaborate more on the data collected and assess some more parameters, and add qualitative assessments based on the information gathered.

This review will take into account additional project/programme documents for the selected interventions and, as appropriate, phone interviews or a web based survey with selected agency desk or technical staff familiar with the projects/programmes. The workload may amount to about 1-2 days per intervention.

The desk-based review will include assessment of actual progress and results at different levels based on indicators reported during implementation or, if available, end reports, evaluations or other available evidence. It will look for evidence of the achievement of results as well as unexpected/unplanned positive and negative effects or distortions. Based on the documentation, and as relevant complemented through interviews, the desk-based review would also involve an assessment of the technical evaluability of the intervention, as it will form the basis for a recommendation for which countries to select for country studies.

Interventions for the desk-based review sample will be selected among the above portfolio sample, maintaining a reasonable level of spread. The selected cases may constitute about 15 interventions (depending on complexity and expected workload) that satisfy the following:

- Availability of progress reports/reviews/final report for at least a 3-year period
- Distinctive feature of the intervention are expected to illuminate (positively or negatively) at least one of the focus areas listed in the ToR or other aspects found of particular interest to this evaluation.
- The sample should include some ‘high-evaluability’ interventions, where one can expect that sufficient data is available and the degree of complexity is manageable, in addition to interventions that may be less easy to evaluate.

Both the selection of the portfolio sample and the zooming-in process on cases for desk-based review are deliberately phrased in relatively soft terms, to allow for taking new features of the interventions into account that cannot be foreseen at this stage.

The desk-based review will be reported in the same format (database entries and scorecards) as the portfolio screening, with more elaboration and more qualitative assessments, as well as a maximum two-page narrative report of key findings for each intervention. These will present the preliminary findings that the country studies should validate and expand on. Based on the desk-based study, the evaluation team will suggest countries to select for country studies.

4. Data to be collected in portfolio screening and desk-based review

Key parameters of the information that will be looked for are listed below. They will undergo modifications early in the inception phase and will then be translated into multiple detailed specific statements/questions (20-50 entries per interventions) to be responded to by the evaluation team. Some of the information will be collected from all interventions during the portfolio screening, while other parts are expected from the desk-based review only. Some of it will be quantifiable, other parts not.

a. *Basic facts*

Title, country, time period, sector, whether it is a programme aiming specifically at capacity development or a component of a larger scheme, involvement of other donors, document sources used for the screening.

b. *Scope and focus*

Characteristics of the partner institution(s) whose capacity development will be supported, the focus areas/objectives of the capacity development efforts at performance level (e.g. quantitative improvements or qualitative changes in service delivery) and at capacity level (e.g. technical skills, management, strategic planning, outreach, technology and infrastructure, etc.), how internal and/or external factors and actors are addressed.

c. *Pre-implementation analysis, theories of change and complexity*

Level and depth of analysis (internal and contextual factors and actors, risks etc.), explicit or implicit articulation of the theories of change (strategies). Apparent complexity of the capacity to be developed. Whether and how issues like incentives, elite interests, social and cultural norms and other context specific but perhaps sensitive issues are addressed.

d. *Level of specification of processes and results*

Characteristics of specified results and indicators, the level of articulation in results management (e.g. in terms of processes, capacity, performance, outcomes for users, wider impact) and the level of specificity of indicators ('SMARTness'). Indicated flexibility and adaptability of the results management approach, articulation of a change management informed approach.

e. *Donor and partner inputs*

Specification of all inputs to capacity development efforts categorized according to type of inputs, quantified in terms of cost as share of the relevant institution's overall turnover.

f. Roles of donors vs partners

Evidence of leadership/participation of donors vs partners in articulation of capacity development needs, CD processes and support to these processes. The roles and responsibilities of leadership and management as well as donors and possible implementation units, and the roles and responsibilities for monitoring, leading to an assessment of degree of 'ownership'. Available information about who took the first initiatives, who wrote the programme document/strategy (donor, partner or external consultants). Assessment of whether the capacity development objectives can be seen as truly shared aspirations between donors and partners or perhaps rather reflect a donor-driven agenda.

g. Evidence of results and effectiveness

All available evidence of results and effectiveness of the support to capacity development, including monitoring data on the achievement of objectives, any unplanned or unintended (positive or negative) effects, available evidence of actual changes in the capacity of the institutions supported (whether likely results of CD efforts or not). This will be based on monitoring data, external reviews and evaluations, and other sources.

h. Assessments: Complexity of capacity; adequacy of CD approach; responsiveness and legitimacy of donor; results-orientation and specification; innovative approaches

The evaluators' narrative assessment of the intervention, including their understanding of the initiation, why it got the shape it got, and how it developed. Assessment of the apparent complexity; the adequacy of the CD approach; the responsiveness of the donor agency; the legitimacy of the donor's role and inputs; the overall fit between the CD endeavors and the development objectives they were meant to serve; and how the results management approach fit to the CD processes. Any approaches or elements of the CD support, which appear innovative or of interest to the further evaluation. If data are based on interviews, particularly interesting testimonials of informants.

5. Country study design

Three countries will be selected for country studies based on information from the above two steps and recommendations from the evaluation team. The evaluation team will suggest country studies in the inception phase, while the final decision rests with the client. Upon selection of countries, all interventions supported by the agency in that country that fall into the scope defined in the Terms of Reference are eligible for further study. If there are many interventions, a further selection of cases for in-depth study may be made within the country. If the interventions selected were not part of the portfolio screening and desk-based review, they will be before the country study is conducted. Comparing the desk-screenings with the findings during the country studies may itself be of interest to check the robustness of screenings based on documents only.

The evaluation team will propose a methodological approach for the country studies, guided by the approach paper (annex 1) and the review of evaluation methodologies prepared as background for this evaluation², for approval by the client.

² Available at www.sida.se/English/About-us/How-we-operate/Sida-Evaluation/Ongoing-evaluations/Capacity-development