

SECTION 2: PROPOSED METHODOLOGY

THE CD PROCESS

The narrative

In accordance with the key issues outlined at the end of Section 2 (paragraph 3.4), the basic narrative of the theory of change, on which the proposed methodology relies, may be expressed as follows.

A significant and sustainable change in the capacity of a given institution (or institutional system), which enables that institution or system to improve its efficiency and effectiveness in the accomplishment of its own mission, is the result of a deep endogenous learning process including:

1. the acquisition of individual and organisational capabilities, *and*
2. their mainstreaming and transformation into an overall institutional capacity encompassing a coherent improvement in a number of basic features summarised below, which need to be adapted to the specific nature of the institution or system and to the characteristics of the context:
 - policy initiative and autonomy,
 - links to the results,
 - institutional networking,
 - flexibility and adaptation, *and*
 - the coherent expression of all such features.

Such change, like the endogenous process that determines it, is made possible by an enabling environment, which drives the change process through the provision of adequate opportunities, visions and resources. The political and economic opportunities that drive the change are provided by the international environment and partnerships and the domestic political leadership. The specific resources to support the change are provided by possible external and internal support programmes, which may have implicit or explicit capacity development components.

The intervention logic

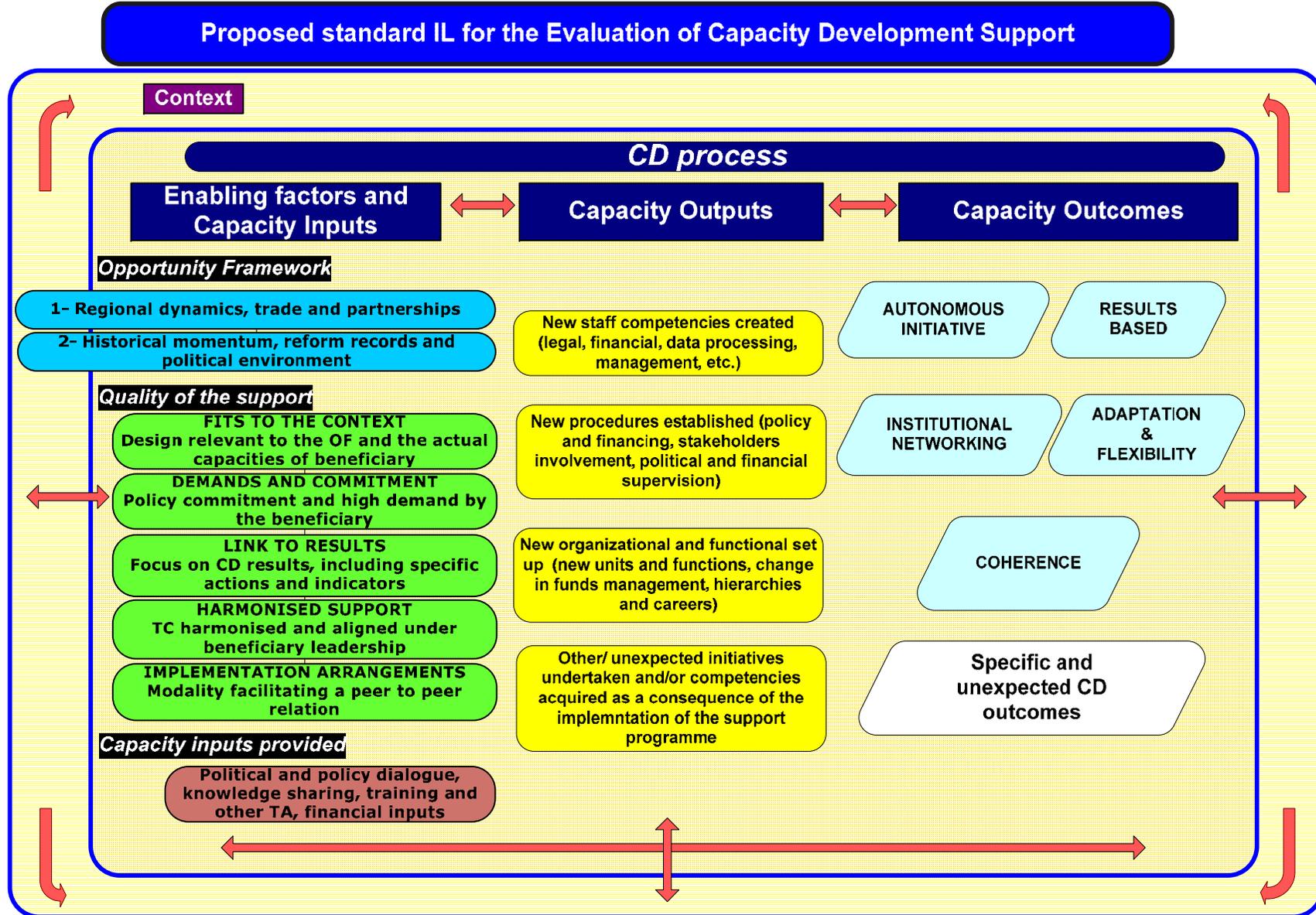
In accordance with the basic EC planning and evaluation methodologies, it has been agreed that the logical framework be used to represent the change process and structure the evaluation, involving the construction of an Intervention Logic (IL) and the identification of a chain of effects linking context, inputs, outputs, outcomes and so forth. It should be made clear that this choice does not affect the actual content of the proposed methodology and other approaches might also be used¹.

The proposed IL shows only the crucial levels (enabling factors and inputs / outputs / outcomes), which may be complemented by other intermediate or longer-term levels (e.g. immediate effects/ induced outputs/ impacts, etc.) according to the depth of the evaluation. This basic IL is shown in Figure 5 below. It describes, in graphic format, the Intervention Logic used in conjunction with the proposed methodology. The IL can be used to support the evaluation methodology guidelines promulgated by the EC-DEVCO Joint Evaluation Unit for “evaluations at programme, project or cross-cutting levels”. It represents a model of capacity change based on a number of concepts of change dynamics that are explained in the Inception Report (including open systems, knowledge reinforcement, and the effects of ownership and leadership on motivation and behavioural dynamics at institutional levels).

¹ Using other instruments, one could emphasise the analysis of the nature and depth of the changes, and give less importance to the causality links, but the assessment of the change processes and its steps would remain the same.

Based on the detailed research carried out under the banner of this mandate, it is clear that evaluators must contextualise the diagram, including basing the logic they propose on local facts and conditions. The IL diagram and its foundational tenets are not meant to represent a generalised model, but rather a CD road-map that needs to be adapted specifically to the issue at hand.

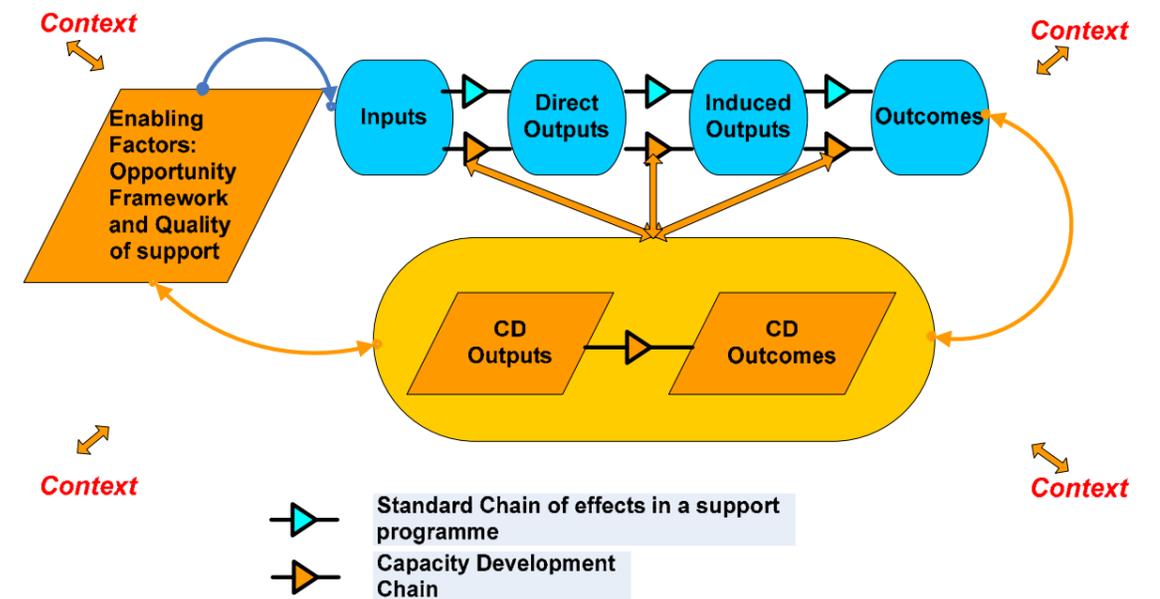
FIGURE 5: PROPOSED STANDARD II FOR THE EVALUATION OF CD



The IL of a capacity development action is usually nested² in the IL of a standard support programme, as suggested in the Figure below in which the implicit CD process is unpacked. The figure shows that the inputs and outputs of a support programme contribute to a capacity development process, with the latter in turn contributing to generation of the effects of the programme (namely the induced outputs and outcomes). The figure, however, also shows other features:

- the IL of the support programme describes two flows of effects: the blue one emphasises the standard sequence of the chain of effects, which does not yet explain to what extent the endogenous capacities have contributed to the determination of the effects. Especially in a short-term perspective, these might have been obtained only (or mainly) through the action of the external Technical Assistance. The brown flow emphasises the contribution of the capacity development process to the determination of effects.
- the Enabling Factors influence both chains of effects (support programme and CD process), although they are determinant for the CD chain, while – at least in theory - the operational chain could function even if they are weak or absent (when external TC is substituted for internal capacity).
- the CD process contributes to the chain of effects of the support programme, but is also affected by its results, by way of the loops shown in the figure.

FIGURE 6: NESTED ILS



There now follows a brief description of the key elements of the proposed Intervention Logic Diagram (refer to Figure 5).

Level 1 of the IL: enabling factors and CD inputs

Level 1 of the proposed IL (Figure 5) contains the Enabling Factors of a CD process, which act as both preconditions for, and key inputs into, the process to take place. These include three different groups of items that affect the entire chain of effects described in the IL:

² The nesting concept is common in the evaluation literature. In our particular case, it is implicit in the ROACH and WBI approaches, while the need to unpack the CD process and identify specific CD outputs and outcomes (within a nesting concept) is explained and developed in the Inception Report of the present study.

- a. The Opportunity Framework (OF), which includes features of the context that in general cannot be influenced by an externally-provided support programme. To a certain extent, however, under certain conditions the OF may be affected by significant partnership arrangements, including political dialogue and the related economic and institutional opportunities. The OF includes two combined dimensions:
- i) first, the momentum of the country in a given phase of its development process. This is the real engine of growth and development, and affects the opportunities and motivations of the institution - or system - that is the subject of the CD evaluation. Within such a framework a TC support programme should be tailored to play a facilitation role. The OF/1 includes such vectors as the historical momentum³; the regional context and related integration⁴; and the specific comprehensive partnership agreements⁵.
 - ii) second, the reform commitment of the government and the political economy that affects the institution - or system - involved. The OF/2 includes the recent political records of change, and the socio-political context that supports it.
- The assessment of the OF should tell if and to what extent the external conditions for the (explicitly or implicitly) intended capacity development are there and what should be done to enhance their conduciveness or to better adapt the support programmes to their actual potential.
- b. The Quality Criteria (QC), that is the quality of the support provided, the way it is conceived, appropriated and implemented; *and*
- c. The actual Support Inputs provided.

The IL considers the inputs that provide the resources for CD from a double point of view:

- from the point of view of the design, appropriation and delivery methods which are quality controlled by the EC through its QSG processes. The Quality Criteria scrutiny should tell if and to what extent the inputs of the support programme (including their design, quality and delivery methods) fit and support the Capacity Development process, so that their high level, in combination with a positive OF, should ensure the attainment of significant capacity outputs and outcomes. The QC now used as the basis for QSG and for part of the ROM processes constitute a strong baseline that can be used to oversee the overall implementation of TC-Reform. There are minor elements of integration into the QC that arise from this methodology. One, for example, addresses the incorporation of strategic institutional contexts into the design of TC and CD, including M&E.
- from the point of view of the specific CD inputs, when they are explicit, including: (i) the political and policy dialogue, which affects or interacts with the OF; (ii) possible knowledge-sharing initiatives, such as inter-institutional exchanges, with regional or international sister institutions, peer-to-peer approaches or twinning experiences; (iii) various types of training;

³ E.g.: Rwanda experiences a new political unity and determination emerging from a deep crisis (rebound effect); Ghana experienced a consolidated history of good governance and growth; Zambia combined a long period of high export prices with an important trade and cooperation partnership with China; Ukraine planned support for Europeanization of agricultural policies in 2007, but when the programme started (in 2010) the country developed opposing policy priorities; Bolivia policies to support coca producers reflected strong political commitment on the part of the new government, where the best energies are invested; etc.

⁴ e.g. the country is included in a fast-growing regional context, the outcomes of which are maximized through specific free trade agreements (e.g.: Vietnam).

⁵ e.g. the case with some ENPI countries, such as Tunisia and Morocco, which have tailored most of their reform processes in the last fifteen years to the integration process with the EU. The impact that has resulted from the expectations of some African countries concerning the establishment of an EPA with the EU is another example. The same applies to different models of partnership (e.g. delivery of commodities against provision of investment), such as those promoted by China with some developing countries.

(iv) different types of TA; and (v) possible financial support to ease the institution's mission and operations.

The QC (according to the headings already adopted by the QSG) are:

- i) Fits to the context. This includes the relevance of the programme *in relation to* the OF and the existing capacities of the beneficiary. *Difference from the present QC: more emphasis on OF.*
- ii) Demands and commitment. This includes the level of policy commitment of the beneficiaries at various levels (e.g. government, specific beneficiary institutions) involved in the sector or themes addressed by the support; and the actual demand for and ownership of the content of the programme. *Difference from the present QC: more emphasis on policy commitment.*
- iii) Harmonised support. This includes the establishment and consolidation of a dialogue framework on the content of the programme driven by the beneficiary and in which other donors participate. The adoption of joint mechanisms, consultation among donors, possible complementarities and other strategic design factors should also be considered. *Difference from the present QC: more emphasis on "sectoral approach".*
- iv) Link to results and expected outcomes. This includes consideration by the programme of specific CD effects in terms of both outputs and outcomes, with specific indicators. *Difference from the present QC: focus on CD results, not only on programme results.*
- v) Implementation arrangements. This includes the TC supply modality and addresses the decision-making process (who manages the programme - a PIU or the beneficiary?), and how the TC is delivered (through a peer-to-peer approach, a traditional consultant-based support approach, or another...). *Difference from the present QC: more emphasis on peer-to-peer (inter-institutional) cooperation.*

This methodology ensures that such enabling factors (both the OF and the QCa) are very well examined. Whereas they are most often relegated to the backdrop within existing evaluation models, they must be well understood in this model because they condition the success of the process affecting the motivation and opportunities for change; they also define the M&E oversight and responses that will be, or have been, applied to CD initiatives. Since this model assumes that constant or 'developmental' evaluation approaches will be applied throughout the life cycle of TC, understanding these vectors is not only important but critical.

Level 2 of the IL: capacity outputs

These are the actual changes in the internal competences and skills that are found in the beneficiary institution(s); they may be directly determined, induced, facilitated or hampered by the implementation of a given support programme. Such outputs do not represent new capabilities *per se*, but identify areas where institutional competence is likely to have been increased through the contribution of the support programme or other resources available in the context.

The changes in competences may be reflected in staff, procedures, knowledge and structures of an institution or system:

- when associated with specific support actions, they appear as direct outputs (e.g. staff trained).
- when conceived as a second-order (indirect) consequence of the support's implementation, they are considered as induced outputs (e.g. new functions that can be fulfilled by the upgraded staff without the benefit of additional CD inputs or outputs).
- finally, there may be cases where such competences are acquired through inputs not directly related to specific support actions, but available in or provided by the context. The IL also makes it possible to capture and assess such competences.

Given the need to keep this methodology within the limits of a relatively simple framework, splitting "outputs" into two parts (i.e. first and second order effects, more simply described as "direct" and "induced") is not required as in other evaluation methodologies. In the event that an evaluation mandate covers a complex institution within a socially or politically complex environment, it is recommended that the evaluation team takes into consideration this difference by focusing on the

induced outputs - which contain greater value-added than direct outputs - while addressing the direct outputs as a lower level of effects.

The IL identifies four categories of output that may be categorised as:

- a. Staff: new staff with new expertise, or new competences among existing staff, with a view to responding better to the institution's mission, may have been the consequence of various actions promoted or facilitated by the programme. Such actions may have included staff recruitment, training and upgrading, exchange of experience, and so forth. The new expertise and competences acquired should enable the institution to fulfil new functions or improve the existing functions (e.g. production of legal and regulatory documents, financial reports, statistical and monitoring reports, etc.).
- b. Procedures: a support programme through its CD component may have contributed to changing and standardising some strategic procedures of the institution, for instance the introduction of systematic stakeholder consultations or the introduction of an MTEF.
- c. Structures: changes in institutional structures, possibly promoted or facilitated by the programme, range from the creation of new units, for example monitoring and evaluation, to the reduction of organisational overlapping, the adoption of a decentralized structure, and so on.
- d. Unexpected: these outputs include other factual changes in the institutional framework (initiatives, responsibilities, competences), which were not planned by the support programme as such but occurred during its implementation and may or may not be placed in relation to such implementation.

If the evaluation team and the developing partner decide to add other categories for one reason or another, this can be accommodated within the boundaries of the methodology. The most important issue to evaluate is the extent to which the outputs, direct or induced, have created additional capabilities (see next level in the logical chain) and whether the combination of those capabilities has given rise to increased capacity in the institution. This evaluation "focus" coincides with the guidelines of the JEU in that it prioritises evaluations that focus on outcomes and impacts.

Level 3 of the IL: Capacity Outcomes

These include the acquisition by the beneficiary institution(s) of new levels of capacity. As shown in the IL the relationship between such capacity and the development results targeted by a given support programme is complex and is not accomplished during the life of the support programme:

- On the one hand, such capacity may or may not have been translated into the expected performance (induced outputs and outcomes) of the support programme under evaluation. This is relatively clear when a programme aims at the achievement of general development indicators. For instance, a programme aims at strengthening the ministry of education and improving access to primary school in rural areas. Having a more powerful ministry and more rural children at school in a relatively short term does not mean that the education system has become stronger. The policy and financial autonomy of the institutions involved, their operational capacity, their relationship with the stakeholders and the final users, and their resilience should all be assessed so as to capture the actual strengthening of the institutional system, and so the institutional sustainability of any possible achievement.

It should be stressed that the performance indicators of a support programme may not be used to assess the capacity development process, even if they are specific CD indicators, since in most cases - as might have been the case in the example of education - they refer to the acquisition of capacity outputs (new competences, functions, structures and funds). The CD process must be assessed from within the institution and its system, through outcome indicators that are sufficiently general and flexible to allow an understanding of achievements that were not pre-determined and have occurred during the process itself.

- On the other hand, it must be noted that this capacity is, by definition, absolutely necessary for the accomplishment of the institution's mission beyond the duration of any specific support programme, and is therefore the basis of the institution's sustainability. There must be a

fundamental distinction between the performance indicators of a support programme and the performance indicators related to the strategic institution's mission.

To identify such capacity various alternatives have been considered⁶. In the end it was agreed to capitalise on the recent best-case experience of the Netherlands evaluation unit (IOB) and adopt a similar approach⁷. This choice integrates the 5Cs approach into the proposed evaluation model. The advantage of this choice is that the 5Cs approach has already been widely tested by the Netherlands Cooperation programme and its adoption by the EC may facilitate strong harmonisation within the EU development policy framework. The 5Cs have been incorporated into Figure 5 above (in the Outcomes column). Some minor changes in the definitions were introduced following the field tests, to make them more understandable and adaptable to the specific frameworks.

As mentioned above, the capacity of an institution or system has to be assessed through the consideration of a number of fundamental capabilities, or types of behaviour, or modalities of action, to show that the institution or system is able to fulfil its mission under different conditions on a relatively long period of time. This is why the recent attempts to establish specific approaches for the assessment of CD have converged on identification of some key features, relatively general and flexible, with content that can be adapted to the different policy and institutional contexts:

- the 5Cs methodology proposes four groups of capabilities plus a comprehensive element to establish coherence among them, such as: to survive and act; to adapt and self-renew; to generate development results; to relate;
- the WBI proposes three main capacity outcomes, such as: strengthening stakeholders ownership - that is the demand institutions; strengthening policy efficiency and organisational effectiveness - that is the supply institutions.
- several mission-based approaches, such as those in use for evaluations of institutions with a relatively competitive mission (e.g. universities⁸), identify some basic capacities, for example: to strategize and plan; to mobilize resources; to operate and attain results; to govern human resources; and to learn by doing.

Table 2 shows the correspondence of the definitions used in the different approaches. It is relatively amazing that through a different conceptual framework and diversified priorities, the various approaches considered converge towards a comparable set of areas. This is important for an understanding that the focus should not be on the specific definitions, as they should come from a careful understanding of the contexts. The focus should rather be on the ability of the definitions adopted to identify institutional behaviour and achievements that may guarantee the accomplishment of the institution's mission on a medium-to-long-term horizon under different conditions, including domestic crises and external shocks.

TABLE 2: COMPARISON BETWEEN THE KEY CD OUTCOMES UNDER DIFFERENT CONVERGENT APPROACHES

⁶ These capacities may be identified in different ways, according to the emphasis of the evaluators, either using or maintaining a strong reference to the institution's mission and functions, or else emphasizing the key behaviour of an institution to fit different missions and specialisations. In a first phase, the present study adopted the first approach, i.e. a mission-based approach. To that end, four key capacities were proposed: capacity to strategize and plan; to mobilize resources; to operate and learn by doing; to manage HR and govern.

⁷ See above, the reference to the 5C approach.

⁸ See EUA, '10 year anniversary: Institutional Evaluation Programme', 2004 and ACCJC 'Guide to Evaluating Institutions', 2010.

TABLE 1: COMPARISON BETWEEN THE KEY CD OUTCOMES UNDER DIFFERENT CONVERGENT APPROACHES

5Cs	WBI	UNDP	Institutional evaluation
capability to survive and act	strengthening policy efficiency	stability	strategize and plan
capability to adapt and self-renew	all (see also intermediate outcomes)	adaptability	mobilise resources
capab. to generate development results	strengthening policy effectiveness	performance	govern change and learn
capability to relate	strengthening stakeholders ownership	all	operate and attain results
			all

Legenda: same colour = strong correspondence - all = correspondent features implicitly mainstreamed in all outcomes

The interaction of the key components of the IL

As explained in section 6.1 the hypothesised CD process is the result of the internal dynamics of a given institution or system, subject to two types of stimuli: a) the driving force of the opportunity framework in which the institution is situated (pulling factors); and b) the quality of the specific support programmes provided (pushing factors).

Under such stimuli the CD process occurs through the acquisition of specific competences and skills at individual or organisational level (Capacity Outputs), which may be appropriated by the institution or system, internalised or metabolised and mainstreamed, so as possibly to generate actual institutional capabilities (Capacity Outcomes).

Both the pulling and pushing factors contribute to all levels of the process (see the logical chain in Figure 6):

- The pushing factors may however be more important in the production of the Capacity Outputs. They may help create some competences and skills, even in the absence of specific opportunities and political support, although the latter are at the origin of the availability of the support programmes and are at least necessary for acquiring the related financial and human resources.
- The pulling factors are fundamental to the actual metabolism of the Capacity Outputs. If there are no genuine opportunities for the establishment of a new education policy and institutional system, for instance, the units and staff trained for sectoral PFM, MTEF, and so forth will migrate to other ministries or even abroad, or will rapidly adopt sub-optimal survival strategies to comply with political patronage. But if the opportunities are there (e.g. there is strong political support, funds are made available by the government, the country is on a growth trend, with good partnerships), the competences and skills acquired are transformed into actual initiative and generate a learning process, with a consolidation of the whole institution or system.

THE KEY STEPS OF THE EVALUATION PROCESS

Although the thorough (standard) evaluation is unlikely to take place with any frequency, it is important to develop its methodology, so as to lay down the conceptual framework for any possible simplified or quick application. To ease understanding of the whole evaluation process, we can use the **3-Step** model⁹. The following assessments should be made:

Preliminary assessments, including:

- an assessment of the Opportunity Framework to highlight the political and economic context in which the institution or system operates and the related driving factors; *and*
- an assessment of the Quality Criteria of the possible support programme(s) included in the evaluation¹⁰.

The assessment of the OF will tell the evaluators to what extent the institution is embedded in a conducive environment, and will be used to better understand the causality links in the CD process, in both Steps 1 and 2.

The assessment of the QCa will highlight how the support programme fits both the OF and the internal institutional dynamics to enhance the capacity development process;

STEP 1 will assess how and to what extent the inputs and activities of the support programme have contributed to generation of capacity outputs in the targeted institution or system, how the QC has affected that contribution, and what has been the role of the OF;

STEP 2 will assess the capacity outcomes attained by the targeted institutions in relation to the capacity outputs and other determining - or facilitating or limiting - factors, namely those relating to the OF;

STEP 3 will assess the causality links between the inputs provided by the support programme and the capacity outcomes attained by the targeted institution(s), in relation to the Enabling Factors (the OF and the QCa).

Preliminary assessments: the OF and the QCa

This phase implies the taking stock of all enabling factors, including context-related factors (Opportunity Framework) and Quality Criteria of the support programme(s). The assessment relies on the existing documentation and, according to the depth of the evaluation, specific studies, interviews or focus groups may be used.

An understanding of the Opportunity Framework (see paragraph 6.3) helps explain the levels of ownership, the actual dynamics and the external driving or limiting factors of the institution or

⁹ See the 3-Step approach adopted by the JEU (DEVCO) for the evaluation of Budget Support. According to this approach the causality link between Inputs and Outcomes should be assessed in two different steps: Inputs → Outputs (Step 1), and Outcomes → Outputs (Step 2). In both steps the approach emphasises the role of the contextual factors intervening in the causal relationships as catalyser or independent causes of the effects assessed. This approach should allow the evaluator to overcome the traps of a linear and deterministic relationship between Inputs and Outcomes, which does not exist in reality. The linear approach tends to overlook the complexity of the process and the participation of multiple factors in the determination of the outcomes. In the 3-Step approach the last Step compares the results of the first and second Steps to find out “how and how much” Inputs have – or have not – actually contributed to the determination of Outcomes.

¹⁰ It should be clear that one can decide to evaluate the CD process in an institution (or system) with or without the presence of specific external support programmes. When there are no external support programmes, only the internal actions directly or indirectly aimed at capacity building will be considered as possible inputs into the process. In such a case there is no assessment of the QCa.

system which affect both the production of the capacity outputs and the generation of the capacity outcomes.

On the other hand the assessment of the QCa highlights the means put in place by the support programme to enable the targeted institutions to profit at the highest level of the existing OF throughout the capacity development process.

One of the main challenges of this phase will be the identification of the interaction between the OF and the intended mission of the institution or system, including the related support action and its QCa. This includes: (i) the extent to which the OF provides a conducive framework for the institution or system and the related support action to attain the respective objectives; and (ii) *vice versa*, the extent to which the institution or system and the related support action are enabled or tailored to respond to the OF features and facilitate its positive influence.

TABLE 3: STANDARD EQS FOR THE PRELIMINARY ASSESSMENT

EQ RELATING TO:	STATEMENT OF THE EQ
EQ1: Opportunity Framework./1	To what extent do the country's historical momentum, growth and partnership opportunities, and other existing contextual factors, affect the institutional context of the CD action?
EQ2: Opportunity Framework./2	To what extent do the reform records of the government and the political environment affect the institutional context of the CD action?
EQ3: Quality Criteria	To what extent does the support programme under evaluation respond to the Quality Criteria established by the EC Backbone strategy for Technical Cooperation, including relevance to the context ¹¹ , ownership of targeted institution(s), strategic focus, harmonisation and appropriate delivery modalities?

Table 3 shows the specific EQs relating to this phase of the evaluation. These EQs, like those that will be proposed for the other phases and steps of the evaluation, are meant to be illustrative only and should be modified (added to, amended, eliminated) to reflect the specific contexts and conditions of the actual evaluations. For that reason no indicators are identified, and only an illustrative set of Judgement Criteria is provided in Annex 1.

STEP ONE – on the production of capacity outputs

Step 1 builds on the basic input-output information gathered through monitoring or – in the EC programmes – through QSG annual processes involving the EUD, but also on specific research related to the production of expected or unexpected capacity outputs in the targeted institutions.

The inputs considered here are all those provided by the external support programme and the related activities, regardless of whether or not they have a specific CD purpose. Other internal inputs are also considered. On the other hand the capacity outputs to be considered include both the expected and unexpected capacity outputs generated during the period under evaluation.

For all such outputs, possible causality links with the inputs will be investigated¹². The role of the OF and the importance of the QC, in the production of the outputs, has to be assessed. Some examples may better explain this relationship:

¹¹ The enlargement of the idea of context to the notion of OF implies an adaptation of the standard QCa. Of particular importance are the political and policy dialogue associated to the support programme, to enable a close interaction with the OF, namely with respect to sectoral policies and/or comprehensive partnerships.

¹² This evaluation methodology uses the same methods as most evaluation methodologies to assess and validate the causality links: i.e. building simple counterfactuals in the Step 1 (including before/after and with/without comparisons, based on informed advices), and also using different quantitative methods in Step 2, according to the complexity of the evaluation.

- some capacity outputs (e.g. creation of new structures and functions in a ministry) may be the direct consequence of the government reform process (OF), without any specific contribution from the inputs and activities of the support programme under evaluation;
- some training offered by a support programme may create stronger skills if accompanied by peer-to-peer exchange of experiences (QC), than it would create if based on traditional professorial teaching (QC);
- some training may produce individual skills that push the beneficiaries to migrate toward other institutions or even abroad, if the institution is not supported politically and its opportunities for growth are limited, while they may be translated into new institutional structures and procedures if the OF is conducive.

The following table contains a list of possible EQs for Step One.

TABLE 4: STANDARD EQS FOR THE STEP 1 ASSESSMENT

EQ4: Outputs- staff competences	To what extent did the programme or other inputs contribute to production of objectively verifiable changes in staff competences (legal, financial, data processing, management...)? How did external factors affect such changes?
EQ5: Outputs- procedures and functions	To what extent did the programme or other inputs contribute to production of objectively verifiable changes in institutional procedures and functions (policy and financing, stakeholders' involvement, accountability and supervision)? How did external factors affect such changes?
EQ6: Outputs- organisations	To what extent did the programme or other inputs contribute to the production of objectively verifiable changes in the organisational and internal functioning (institutional structure, decision process, internal mobility and competition)? How did external factors affect such changes?
EQ7: Outputs- unexpected	To what extent did the programme or other inputs contribute to the production of objectively verifiable changes in respect of individuals, organisations and initiatives that were not targeted? How did external factors affect such changes?

STEP TWO - on the emergence of CD outcomes

The scope of this step is to assess the actual changes in CD in the targeted institutions, according to the capacity outcomes identified in the IL: initiative, results, networking, adaptation, and coherence. During the three Rapid Assessment tests of the present methodology, the 5Cs – which remain the reference for the capacity outcomes mentioned – have been renamed so as to facilitate their unambiguous identification by the stakeholders involved in the assessments and their adaptation to the specific contexts.

In this step the assessment is also extended to the causal links between the capacity outcomes and the capacity outputs or other factors relating to the Opportunity Framework.

Table 5 is meant to show the key EQs that could be applicable to the outcome level of the methodology proposed. The six EQs in Table 5 may present a formidable amount of research for an evaluation, but it should be recalled that the evaluators adjust them to the specific context and then choose appropriate indicators.

The wording in the EQs within the table is somewhat generic because it is proposed as a means of understanding the evaluation methodology; during an actual evaluation the wording would be adapted to the context and particular attention should be paid to the institutional and organisational environment (including the Opportunity Framework and the policy and reform realities) within which the CD objectives would be set.

TABLE 5: STANDARD EQS FOR EXPECTED AND UNEXPECTED OUTCOMES

EQ8: Initiative	To what extent is the institution more capable of generating plans (at strategic or other levels) that reflect its stated needs, mission and various changing environments and then mobilizing its resources and management to execute them?
-----------------	--

EQ9: Results	To what extent is the institution more capable of achieving and monitoring the “developmental results” stated in national and “departmental” plans in a sustainable manner?
EQ10: Networking	To what extent is the institution accountable and able to work in a coordinated and efficient manner as part of a wider network of interested stakeholders?
EQ11: Adaptation	To what extent is the institution in a position to adapt constantly in response to changing external environments and conditions?
EQ12: Coherence	To what extent has the institution succeeded in putting in place policy and management frameworks that build on one another and provide evidence of a clear chain of results from the strategic to the operational levels?
EQ13: Unexpected outcomes	How have non-planned and/or context-specific capabilities (developed as a result of Capacity Development efforts in the institution) improved or reduced the overall capacity of the institution to carry out its vision and achieve its objectives?
EQ14 to 17: Causality links	To what extent have the institutional capacity outputs and/or other factors related to the OF contributed to each of the above-mentioned capacity outcomes (initiative, results, networking and adaptation).

STEP THREE: Causality links between the CD inputs and the CD outcomes.

According to the 3-Step approach it is difficult to use a linear model to assess the direct link between the inputs provided and the outcomes generated. This is particularly true in our case, as the process that leads to the capacity outcomes is complex and implies the contribution of so many factors, namely the OF and other institutional dynamics. The causality link between the CD inputs and the CD outcomes (STEP 3) has to be assessed through a systematic comparison of the results of Steps 1 and 2. Step 2 shows how changes in competences and experience have or have not contributed to an increase in capacity outcomes, in the framework of a given context. Step 1 shows how the programme inputs have had any influence on such competences and experience, again in the framework of a given context. Step 3 highlights the transitive relationship between inputs and outcomes. Formulating specific EQs for such an assessment is unnecessary¹³.

Link between CD evaluation and standard programme evaluation

CD and standard evaluations are not superimposable

A clear distinction should be made, in the short term, between the evaluation of an institutional CD process and the evaluation of the performance of the same institution *vis-à-vis* a set of externally given objectives, as is the case when evaluating a development programme.

The CD evaluation aims at identifying the progress achieved, within the institution, in terms of skills, competences, strategic initiative, implementation capacity, and so forth, with a view to long-term fulfilment of the institution’s mission. The standard programme evaluation aims at identifying the progress achieved, during the life of the programme, towards fulfilment of a set of objectives and performance indicators that are coherent with the institution’s mission.

The CD evaluation assesses the strengthening of an institution or system, while the programme evaluation assesses the strengthening of its performance. The two approaches may not be superimposable in the short or even medium term, while they should be so in the longer term provided there is actual correspondence between institutional mission and planned performance (see also paragraph 6.5).

¹³ The three-Step approach has been positively tested in several multi-donor Budget Support evaluations led by the EC DEVCO Evaluation Unit (Tunisia, Mali and Zambia).

It may also be difficult to attempt to carry out the two evaluations in parallel; their objectives may conflict. In particular, in a standard programme evaluation it may happen that the institutions involved feel they are under examination. This may cause a defensive attitude and jeopardise their collaboration in the CD evaluation.

The need for complementarity

The above considerations, however, should not lead to the conclusion that the two assessments should be completely separate. Indeed their complementarity appears ever more important. In particular a standard programme evaluation would benefit much from the availability of an updated CD evaluation of the main institutions involved in the programme. The CD evaluation would improve understanding of the reasons for the successes and failures of the programme, and would allow an in-depth assessment of the sustainability of its results.

The key value-added imparted by a CD evaluation to a standard evaluation concerns the assessment of the **sustainability** of the induced outputs and outcomes. Various cases may arise, viz.:

- Both the standard and the CD evaluations give compatible positive or negative results: this means that the induced policy outputs and the related outcomes of the standard programme are either both positive and institutionally sustainable, or negative and institutionally unamendable.
- The CD evaluation is positive, while the standard programme evaluation is negative: this implies a question of time. The new capacities are not yet translated into new induced outputs and development outcomes, or else they were badly formulated.
- The standard programme evaluation is positive, while the CD evaluation is negative: the induced outputs and development outcomes are not likely to be sustainable. This is for instance the case in many countries where intensive TA programmes are implemented.

To make CD evaluations available as a key complement of the evaluation process is a complex issue, if excessive organisational burden and duplication is to be avoided. It is recommended that part of the CD evaluations be integrated into the recurrent assessments carried out by the EUDs and the monitoring system - that is the preliminary steps - and that methods of rapid assessment to carry out CD evaluations be identified a few months prior to the planned programme evaluations for selected programmes.

Conditions for carrying out a rapid CD assessment

Provided that sound quick assessment tools are available, as proposed in the following chapters of this report, the complementarity between CD evaluation and standard evaluations may be ensured on a systematic basis.

CD evaluations should be carried out on all programmes with a significant TC component, including the first three categories of TC identified by the Backbone strategy (capacity development, policy advice, support to service delivery).

How can one establish whether a TC component is significant or not? Several criteria should be used to determine whether the following apply:

- in the case of standard TC programmes, when a programme supports the establishment of a sectoral or thematic approach, including policy and institutional change, with a focus on specific partner institutions or institutional systems (e.g. at sectoral and local levels). There should also be a particular level of TC (say above €400,000 per year);
- in the case of Budget Support programmes, where financial resources are provided to specific partner institutions (or institutional systems) at country, regional, sectoral level, to strengthen their effectiveness on sustainable bases, with or without specific TC components;
- in the case of support to civil society via NGOs and other Non-State Actors, provided that the programmes have a relatively wide scope and have a well-defined partnership with specific institutional systems;
- finally it would not be advisable to carry out a rapid CD assessment of a comprehensive country or regional programme as such, as it would be difficult to identify the right institutional dimension.

Planning the rapid CD assessment could be either independent or combined with standard evaluations. The EUDs should decide each year the programmes for which a rapid CD assessment would be necessary. At the same time, when a final evaluation of an important programme (sectoral policy, budget, or civil society support) has to be carried out, it would be opportune to plan a rapid CD assessment between six and three months before the evaluation starts.

Besides such planning criteria, the rapid CD assessment should be a flexible instrument, to use on demand. For lengthy programmes (say more than four years), the CD assessment could be repeated twice (mid-term and final). For the types of programmes mentioned above (TC, BS and support to Civil Society), a form of Rapid CD assessment of the beneficiary institutions should be incorporated in the appraisal phase. In such case, the relevant inputs should be those which exist in the institution and in the specific context, before the support programme starts.