
# DG NEAR

# Guidelines on linking planning/programming, monitoring and evaluation

July 2016

Neighbourhood and Enlargement Negotiations

# PREFACE

Managing for results has become a political priority for the Commission. Our institution needs to demonstrate the concrete impact not only of its financial cooperation but more widely of all its interventions. In addition, budgetary constraints, the increasing demand of the general public for the Commission to ensure greater accountability, as well as the need to draw lessons from past experience, all require a greater attention to ensuring impact.

For DG NEAR, which is confronted with a challenging policy context in both the enlargement and neighbourhood regions, ensuring and demonstrating the clear impact of its interventions is a particularly pressing policy priority.

In such a context, it is indispensable to carefully design new policies and actions, to monitor implementation through relevant indicators, identify the main results and determine the EU contribution to these results. We need to ensure that our policy decision-making processes as well as our planning and programming exercises are fed by relevant data, knowledge and lessons learned.

All this requires a strong ability to design interventions, as well as a deep understanding of the links between planning/programming, implementation (and its monitoring) and evaluation.

Therefore, I am pleased that our DG has now prepared these new guidelines. They have been finalised after careful consultations with many staff members across all Directorates of our DG and Delegations, as well as with the Commission's Secretariat General and DG DEVCO.

They reflect DG NEAR specificities and are also aligned with the Commission’s better regulation priorities. They will remain a living document which can be further improved as our DG further develops.

I am confident that these Guidelines will help you and also our beneficiaries, to implement our ambitions in a more effective and efficient way.

*Christian Danielsson*

# TABLE OF CONTENTS

## LIST OF FIGURES



## LIST OF TABLES



## LIST OF BOXES

## LIST OF ANNEXES (VOLUME 2):

# LIST OF ACRONYMS

| | |
|---|---|
| CBC | Cross-border cooperation |
| CIR | Regulation (EU) No 236/2014 laying down common rules and procedures for the implementation of the Union's instruments for financing external action |
| CMO | Context-Mechanism-Outcome |
| CoTE | DG NEAR Centres of Thematic expertise |
| CREAM | Clear, relevant, economic, adequate, monitorable |
| CRIS | Common RELEX Information System |
| CSO | Civil Society Organisation |
| CSP | Country Strategy paper |
| DG | Directorate General |
| DG BUDG | Directorate General for Budget |
| DG DEVCO | Directorate General Development Cooperation and International Development |
| DG NEAR | Directorate General Neighbourhood and Enlargement Negotiations |
| EAMR | External Assistance Management report |
| EC | European Commission |
| EEAS | European External Action service |
| EIF | European Integration Facility |
| ENI | European Neighbourhood Instrument |
| EQ | Evaluation Question |
| EU | European Union |
| EUD | European Union Delegation |
| HQ | Headquarters |
| ICT | Information and Communication System |
| IPA | Instrument for Pre-accession Assistance |
| ISG | Inter-Service Steering Group |
| JC | Judgment criteria |
| M&E | Monitoring and Evaluation |
| MCSP | Multi-Country Strategy Paper |
| MIS | DG NEAR Management Information System |
| NAO | National Authorising Office |

| | |
|---|---|
| NGO | Non-Governmental Organisation |
| NIPAC | National IPA Coordinator |
| OECD–DAC | Organisation for Economic Cooperation and Development-Development Assistance Committee |
| OSC | On the spot check |
| RACER | Relevant, accepted, credible, easy to monitor, robust |
| RG | Reference Group |
| ROM | Results oriented monitoring |
| SDGs | Sustainable Development Goals |
| SG | Secretary General of the European Commission |
| SMART | Specific, measurable, attainable/achievable, relevant, time bound or significant, measurable, appropriate, resourced, timely available |
| SME | Small and Medium enterprise |
| SSF | Single Support Framework |
| SWD | EC Staff Working Document |
| TFEU | Treaty on the Functioning of the European Union |
| ToR | Terms of Reference |

# GLOSSARY

| | |
|---|---|
| Action | A coherent set of co-ordinated activities undertaken to meet a defined objective of a geographic and/or sectorial scope, which have an estimated total cost to which the EU approves a maximum contribution, as well as an implementation schedule and performance parameters. It can be used to refer indifferently to the concept of project or programme.<br><br>The use of the term 'Action' provided in the Regulation (EU) No 236/2014 laying down common rules and procedures for the implementation of the Union's instruments for financing external action is consistent with this definition, even though a certain ambiguity exists in the same Regulation as sometimes action is referred to in addition to projects and programmes.<br><br>For the purpose of the present guidelines though, 'Action' will be used to refer indifferently to the concept of project or programme. |
| Action document | The document providing details on the Action to be funded by ENI and IPA II. Under IPA II Action Documents are used for reference and for information purposes; whereas they form part of the Commission Implementing Decision (i.e. Financing Decisions) for ENI. |
| Action programme | The level at which Financing Decisions are set according to the Common Implementing Rules for External Actions (CiR). The Action Programme is the annex to the Financing Decision. It is a synthesis of the specific Action Documents. |
| Baseline | The value assumed by a given indicator at time $t_0$, against which progress will be assessed. |
| Composite indicator | Formed when individual indicators are compiled into a single index, on the basis of an underlying model of the multi-dimensional concept that is being measured. (OECD) |
| Effectiveness | A measure of the extent to which the intervention's objectives were achieved, or are expected to be achieved (OECD) |
| Efficiency | Efficiency considers the relationship between the resources used by an intervention and the changes it generates (which may be positive or negative). (Better Regulation) |
| Evaluation | The systematic and objective assessment of an on-going or completed action or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability. (OECD) The Better Regulation package defines evaluation as the assessment of the effectiveness, efficiency, coherence, relevance and EU added value of one single EU intervention, thus adding coherence and EU value added.<br><br>An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision– making process of both recipients and donors.<br><br>Evaluation also refers to the process of determining the worth or significance of an intervention. An assessment, as systematic and objective as possible, of a planned, on-going, or completed intervention (OECD). An evaluation can be carried out on various levels: policy, strategy, sector, theme, country, region, programme, project, etc. |
| Goal | The higher-order objective to which an intervention is intended to contribute. (OECD) |
| Impact | Impact relates to the changes that are expected to happen due among other things to the implementation of an intervention. Such impacts may occur over different timescales and affect different actors. They can be positive and negative, direct and indirect, intended or unintended, on any dimension (social, economic, environmental, political, etc. |
| Impact indicator | Signals to which extent the overall (mid and long-term) objectives of an intervention (the strategic objectives or goals) have been achieved. |
| Indicative Strategy Papers (relevant for IPA II) | The Indicative Country Strategy Papers (CSP) and the Multi-Country Strategy Paper (MCSP) are the overarching strategic planning documents from which priorities and objectives of individual programmes derive. They are Implementing Acts (Art. 291 TFEU) adopted by the European Commission following the opinion of the IPA II Committee. |

| | |
|---|---|
| Indicator | A quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect changes connected to an intervention, or to help assess the performance of an actor (OECD). A variable relevant to assess the degree of achievement of a given objective. |
| Input | The political, technical, financial, human, and material resources put in place to generate activities. |
| Input indicator | An indicator that measures the resources and means provided by donors and implementers |
| Intervention | A generic expression referring to the coordinated set of activities and means put in place to implement a given strategy/objective. It can be a project, a complex programme (articulated around a set of projects, a budget support operation or a mix of BS and other typologies of contracts), a policy, a legislation, an action plan, etc. |
| Monitoring | Monitoring is a continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an on-going intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds. (OECD) |
| Multi-annual action programme | In compliance with Regulation No 236/2014, laying down common rules and procedures for the implementation of the Union's instruments for financing external action, the Commission shall adopt multi-annual action programmes. It defines an EU response strategy on the basis of a brief analysis of the situation in the country concerned, and of its relationships with the EU, the relevant partner countries' strategies or plans, the priorities for EU support and the indicative level of funding. They can be drafted for a period of up to three years in the case of recurrent actions, and for a period of up to seven years under IPA II. |
| Objective | The level of achievement expected from the implementation of the goals (OECD) |
| On the spot checks | They refer to a field mission jointly implemented by programme managers and contract and finance staff to assess the status of implementation looking both at operational and contractual and financial matters. |
| Outcome | The likely or achieved short-term and medium-term effects of an intervention's outputs. (OECD) Short to medium term effects on the political, social, economic and environmental areas targeted by ENI/IPA II financed interventions as well as changes in behaviour of addressees of ENI/IPA II financed interventions.<br><br>Other external factors and players also influence the targeted areas and addressee. |
| Outcome indicator | These signal whether the short to mid-term desired changes are happening. |
| Output | Direct products or services delivered by activities, directly influencing the achievement of outcomes. |
| Output indicator | They show the degree of achievement of the direct products of an activity or set of activities. They are by nature activity-specific. |
| Performance indicator | A variable that allows the verification of changes in the intervention or shows results relative to what was planned. (OECD) |
| Process | The modality by which, in a given time frame, inputs, actors and other factors interact during the period of implementation of an intervention leading to the production of the actual outputs and outcomes. It can refer to both the production dimension (how inputs are transformed in outputs, looking also at costs and delays) and the relational side (how actors interact). |
| Process indicator | They measure what happens during implementation and it mainly focuses on the activities execution. |
| Programme | A Programme can have various meanings, either: (i) a set of projects put together under the overall framework of a common Overall Objective/Goal; (ii) an ongoing set of initiatives/services that support common objectives (i.e. a Primary Health Care Programme); or (iii) a Sector Programme, which is defined by the responsible government's sector policy (i.e. a Health Sector Programme), often implemented by means of budget support. A programme may cut across sectors, themes and/or geographic areas. |

| | |
|---|---|
| Project | A Project is a series of activities aimed at bringing about clearly specified objectives within a defined time-period and with a defined budget. This definition allows for great adaptability to the needs of countries and strategies. It can apply to an articulation of activities (services, works, supplies), or to very specific interventions within an Action Document (grant projects funded by a grant scheme; a twinning). In the IPA II language the concept of project is replaced by the concept of action, while in the ENI dimension project is very much used as an alternative to that of a programme, normally, but not only, used for budget support operations. |
| Proxy indicator | Indirect measure or sign that approximates or represents a phenomenon in the absence of a direct measure or sign (Business dictionary). |
| Relevance | Looks at the relationship between the needs and problems in society and the objectives of the intervention. Things change over time - certain objectives may be met or superseded; needs and problems change, new ones arise. The OECD DAC attributes a wider meaning to this criterion, including also the correspondence between instruments and goals—a set of questions which is fundamental in any evaluation and which may overlap with the criterion of internal coherence. |
| Result | The output, outcome or impact (intended or unintended, positive and/or negative) of an (OECD).<br><br>In the present guidelines result will be assimilated to outcome (except when referring to a result's chain) |
| Single Support Framework | In compliance with Regulation No 232/2014, establishing ENI instrument, for countries where jointly agreed Partnership Priorities or equivalent exists, a multi-annual Single Support Framework (SSF) is adopted. The SSF outlines priorities for Union support, selected from those included in Partnership Priorities or equivalent. |
| Sustainability | Sustainability relates to the continuation of benefits from an intervention after major support has been completed. The probability of continued long-term benefits. The resilience to risk of the net benefit flows over time (adapted from the OECD-DAC). It has various dimensions: social, economic, political, environmental, financial, institutional, etc. |
| Target | The specific level that we want to achieve at a given time with regard to a specific indicator |

# 1. INTRODUCTION

This document is a response to the need to strengthen the linkage between planning/programming[1], monitoring and evaluation function in DG NEAR in the context of the neighbourhood and enlargement policies (i.e. the use of the intervention logic and related indicators in monitoring and evaluation exercises, among other reporting requirements), to promote the monitoring and evaluation culture within the DG and to guide the monitoring and evaluation actors in the implementation of the monitoring and evaluation cycles.

The guidelines, therefore, are addressed to:

- DG NEAR geographical and regional programme units, responsible for planning/programming, follow-up on implementation, monitoring and/or evaluation IPA II and ENI;
- DG NEAR monitoring and evaluation officers;
- EU Delegations (EUDs) in the Neighbourhood and Enlargement countries and in particular to the programme managers, monitoring and evaluation focal points, Head of operations, etc.;
- National authorities in ENI and IPA II partner countries – monitoring and evaluation experts at the NIPAC office, NAO office, Operating structures, etc.;
- Implementing partners, other than those mentioned in the previous bullet point;
- External evaluators and monitoring (ROM) contractors.

The guidelines are built on the basis of the requirements, set out in several relevant EC Regulations[2] and related documents (e.g. Framework Agreements), the Better Regulation Package[3], the ROM Guidelines, the IPA II and ENI planning/programming guidelines and methodological notes, issued by DG NEAR[4]. In addition, they take into account the complementarity between monitoring, evaluation and audit functions, as mentioned in the DG NEAR Instructions on preparation of monitoring, evaluation and audit plans[5].

The guidelines do not provide specific instructions on the application of the impact assessment or on the Fitness check[6]. In these cases, the Better Regulation Guidelines are fully applicable.

The document is divided in four parts:

**Part 0** lays down basic definitions and principles of monitoring and evaluation. It presents the uses and the links between these two functions, their complementarity and the scope of their application. Finally, it

[1] Considering that the present document is to be used by stakeholders involved in ENI and IPA II related support, it has been decided using both planning and programming terms. Indeed: i) for ENI programming (preparation of multiannual indicative programmes) comes before planning (preparation of annual/multiannual programmes), and ii) for IPA II: (strategic) planning (preparation of indicative strategy papers) comes before programming (preparation of annual/multiannual programmes).

[2] The Regulations that are referred are specified in the Reference part of this document.

[3] The Better Regulation Package was approved by the European Commission on 19.05.2015 – See SWD (2015) 111 at http://ec.europa.eu/smart-regulation/guidelines/docs/swd_br_guidelines_en.pdf. Please refer to Annex 1 for a synthesis of the main elements of the Better Regulation.

[4] All cited documents are outlined in the Reference part of this document. Please also refer to the C4D website, managed by DEVCO, as complementary source of information and guidance: http://capacity4dev.ec.europa.eu.

[5] See ARES Note n. 3626990 dated 03/09/2015, as far as audit plan for Neighbourhood; ARES Note n. 4235431 dated 12/10/2015 with regard to the risk assessment and audit plan for Enlargement countries; ARES Note n. 4796302 dated 03/11/2015 with reference to the evaluation plan.

[6] A Fitness Check examines available evidence to judge the cumulative results of a group of measures of different nature (spending interventions, regulations, etc.) which share a relationship (e.g., common objectives). Fitness Checks focus on synergies and inefficiencies among the group of measures (European Commission, 2015b).

presents the users, as well as the respective roles and responsibilities of all actors involved in both monitoring and evaluation exercises.

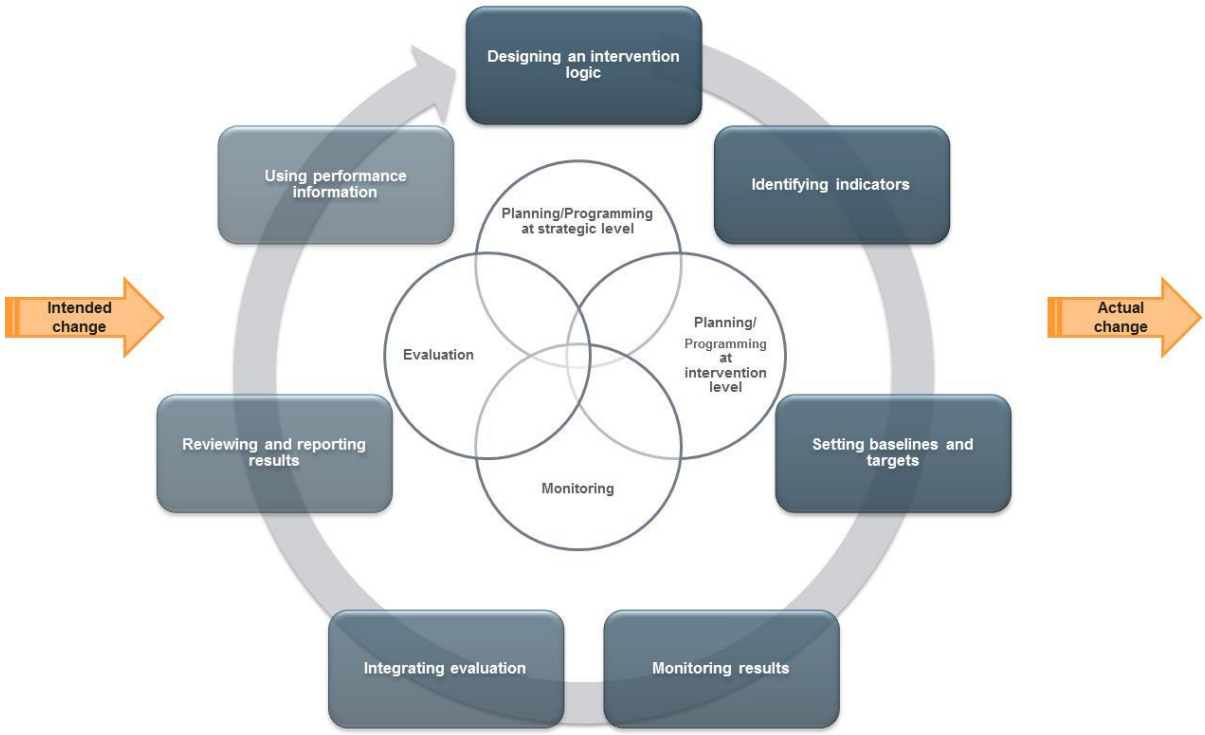
**Part 1** links planning/programming, monitoring and evaluation. It presents key considerations and steps to follow in both the definition of an intervention logic and in that of its corresponding monitoring system. The identification and definition of indicators play a key role in this regard[7].

**Part 2** focuses on the monitoring cycle, from planning to dissemination and follow-up. Even though some aspects are also relevant for external monitoring (ROM), it mainly focuses on internal (both EC/EUD and implementing partners) monitoring.

**Part 3** describes the evaluation cycle, from planning to dissemination and follow-up, providing methodological guidance. It also explores the concepts and significance of evaluation questions in the context of DG NEAR.

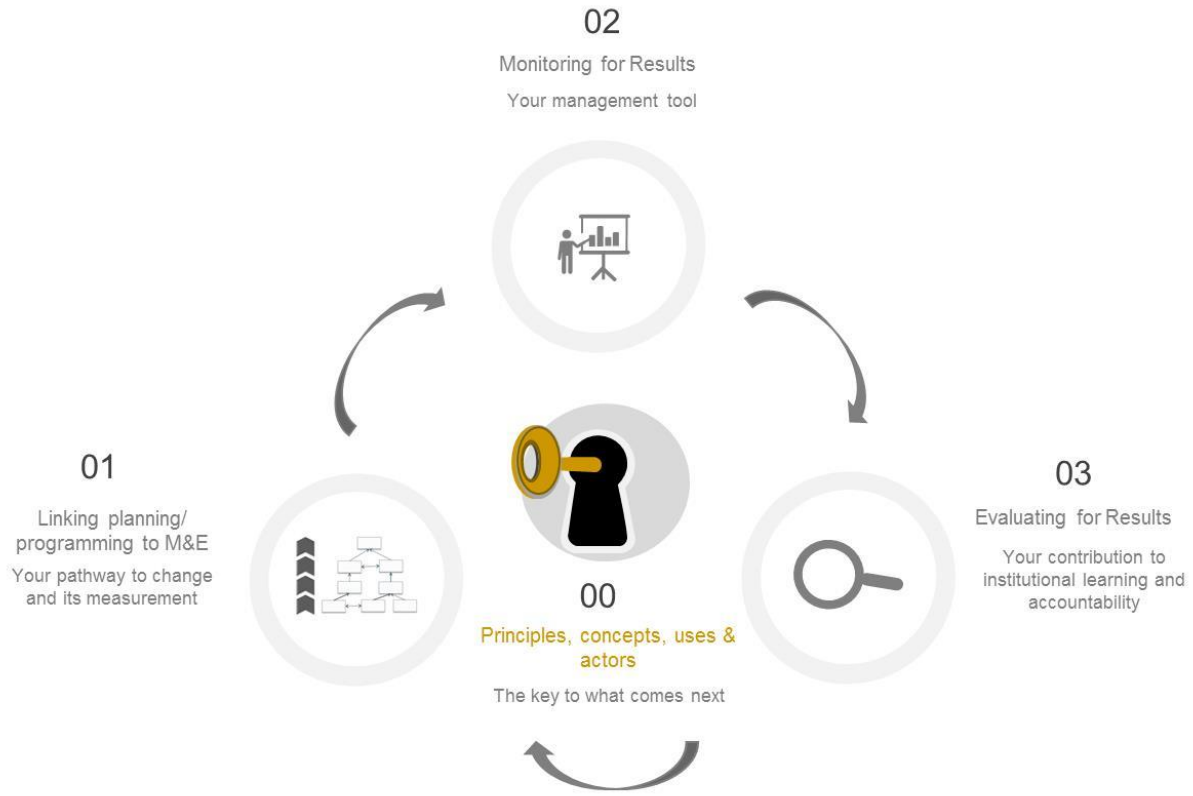
These four parts synthesise the main concepts related to the way in which the intended change sought by means of a strategy and its operational modalities (a policy and a set of actions) can generate the actual change, when the planning of activities, their monitoring and evaluation interact contributing to the final result.

**FIGURE 1: FROM INTENDED TO ACTUAL CHANGE: THE INTERACTION BETWEEN PLANNING/PROGRAMMING, MONITORING AND EVALUATION**




*Source: DG NEAR*

[7] Though these guidelines provide guidance on the way indicators should be designed, they do not explicitly address the way they are used in the framework of budget support operations, for which the interested reader should refer to the Budget Support Guidelines. Refer to: https://myintracomm.ec.europa.eu/dg/devco/eu-development-policy/budget-support-public-finance-management/_layouts/WopiFrame.aspx?sourcedoc=/dg/devco/eu-development-policy/budget-support-public-finance-management/Documents/BS_Guidelines-Part_II_Programming,Design,Management.doc&action=default&DefaultItemOpen=1.




# PART 0: KEY PRINCIPLES AND CONCEPTS, USES AND ACTORS

**BOX 1: SUMMARY OF KEY MESSAGES ADDRESSED IN PART 0**

***Monitoring and evaluation are two separate yet complementary areas of practice providing different inputs into the decision-making process at different points in time.***

***Monitoring aims at answering the following question: What is the intervention doing (from inputs to outcomes) and to what extent its implementation is progressing?***

***While evaluation answers the following one: To what extent and why is the intervention producing (or failing to produce) the specific outcomes and impacts which have actually materialised—be those negative or positive?***

***The Commission Better Regulation Guidelines adopted in 2015 define a set of monitoring principles to be considered when defining a monitoring system: Comprehensive, Proportionate, Minimise overlap, Timeliness, Accessibility.***

***On the other hand, evaluation principles inform the entire evaluation policy, starting from institutional arrangements, as well as the conduct of internal or external evaluators, contracting authorities, and line managers***

***The Better Regulation guidelines define five mandatory evaluation criteria (Relevance, Effectiveness, Efficiency, coherence and EU added value), but other criteria, namely sustainability and impact, should also be addressed in DG NEAR evaluations whenever relevant.***

***Monitoring and evaluation play a key role in the credibility of an institution: they provide among other things for institutional learning and accountability resulting in a better informed governance and thus in a better attainment of EU and partner countries overall objectives.***

***The use of evaluation does not happen automatically. Rather, in order to ensure the use of evaluation findings, evaluation managers have to carefully arrange key moments during the evaluation process.***

***The quality of monitoring and evaluation exercises does not only depend on the quality of their respective deliverables. The action of the different actors involved, each of them having clearly defined roles and responsibilities from the planning until the dissemination and follow-up phases, is equally important.***

*Source: DG NEAR*

# 2. DEFINITIONS

## 2.1. Monitoring

Monitoring is a *continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an on-going intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds*[8].

Monitoring, while being focused on the input, activities, outputs and outcomes levels, aims at measuring progress in relation to the expected outputs, outcomes and impact of an intervention by means of RACER and SMART[9] indicators related to a baseline situation, defined during planning/programming. As such, monitoring is expected to support: i) effective and timely decision-making, ii) learning by interventions' stakeholders, and iii) accountability on the use of resources.

Monitoring can also refer to a regular review of the system put in place to deliver assistance, by the donor agency or by the Government to which management responsibilities have been entrusted[10].

## 2.2. Evaluation

According to the OECD-DAC[11], evaluation is *the systematic and objective assessment of an on-going or completed* intervention[12]*, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability.* The Better Regulation[13] defines evaluation as the *assessment of the effectiveness, efficiency, coherence, relevance and EU added-value of one single EU intervention,* thus adding coherence and EU value added.

The OECD-DAC definition continues as follows: *an evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision– making process of both recipients and donors. Evaluation also refers to the process of determining the worth or significance of an activity, policy or* intervention. *An assessment, as systematic and objective as possible, of a planned, on-going, or completed intervention.*

[8] This is the definition given by the OECD-DAC in its 2002 Glossary of Key Terms in Evaluation and Results Based Management. The definition provided in the Better Regulation Guidelines SWD (2015) 111 final (*Monitoring describes the process of tracking the implementation and application of EU legislation by the Commission or the progress of spending programmes*) is consistent with the OECD-DAC but yet it is considered as being broader. This is the reason why the OECD-DAC's definition is used in the present document.

[9] Relevant, Accepted, Credible, Easy and Robust. And Specific, Measurable, Achievable, Realistic and Time-related. Refer to § **9.2.3 WHAT ARE THE CHARACTERISTICS OF INDICATORS?** for further details.

[10] In the framework of the implementation of a policy, like the Enlargement Strategy or the European Neighbourhood Policy, as far as DG NEAR is concerned, monitoring also refers to the modalities put in place (ad hoc committees, reporting) to review the process of implementation of reforms related to the specific partnership agreement existing among the players (EC and partners countries) within action plans or negotiating chapters. These guidelines do not address this dimension.

[11] Glossary of Key Terms in Evaluation and Results Based Management, OECD-DAC, 2010.

[12] The OECD-DAC refers to project and programme. Nevertheless, since DG NEAR's support goes beyond these, the broader term 'intervention' (that might encompass a coordinated set of activities and means put in place to implement a given strategy/objective. It can be a project, a complex programme (articulated around a set of projects, a budget support operation or a mix of BS and other typologies of contracts), a policy, a legislation, an action plan, etc.) is used in the present document.

[13] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Better regulation for better results - An EU agenda, COM(2015) 215 final. For the Better Regulation Guidelines, refer to http://ec.europa.eu/smart-regulation/guidelines/index_en.htm.

# 3. PRINCIPLES AND STANDARDS

## 3.1. Monitoring principles

As discussed among other under § 8 DEFINING THE INTERVENTION LOGIC AS A STARTING POINT, monitoring is directly linked to the intervention logic. Indeed, **the results' statements included in the intervention logic define what is to be measured, via specific indicators, in monitoring (and evaluation) exercises**.

In this framework, as stated in the Better Regulation, consideration needs to be given to the frequency and method of collection, different sources of evidence, etc. taking into account what is already available and cost for different actors involved. This leads, also according to the Better Regulation, to a series of questions a monitoring system must address:

- What evidence needs to be collected?
- When and how should evidence be collected?
- Who will collect the evidence and from whom?

The answers to these questions are to be governed by a set of principles (refer to BOX 2: MONITORING PRINCIPLES AS DEFINED BY THE BETTER REGULATION).

**BOX 2: MONITORING PRINCIPLES AS DEFINED BY THE BETTER REGULATION**

**Comprehensive:** The monitoring system put in place must cover the objectives of the intervention. Whilst high level monitoring of key indicators linked to general objectives is likely to exist already, new monitoring arrangements may need to be identified for specific or operational objectives. Although monitoring systems generally collect objective (e.g. factual, quantitative) evidence, it is also possible to put in place monitoring of subjective (e.g. opinion based, qualitative) evidence such as periodic opinion polls or surveys.

**Proportionate:** The system put in place needs to reflect the importance placed on different aspects of the intervention. Collection of evidence comes at a cost and care should be taken to challenge the necessity of each (new) monitoring requirement being considered.

**Minimise overlap:** The EU, Member States, Agencies, international organisations collect a lot of evidence. It is important to know what we have already and when and how it is collected. This should help to avoid duplication and the creation of unnecessary data collection burdens by concentrating only on the gaps that need to be filled.

**Timeliness**: Whilst the monitoring system should be set up as soon as possible after the intervention is agreed, this does not mean that all evidence needs to be collected from that point onwards. Not all evidence needs to be collected at the same time – sometimes it is better to collect evidence as it happens, other times it can be done later. Consideration also needs to be given to when the evidence will be used – different legal requirements may play a role not just in deciding what is needed, but also the when it is collected.

**Accessibility:** In principle, all evidence gathered should be made available to the general public, unless data includes confidential elements. Even though this principle relates to data produced by the institutions and other bodies of the European union, and it is linked to the EU Open Data Portal, in the framework of DG NEAR interventions this can be understood as evidence being available to key stakeholders directly affected by interventions' outcomes.

*Source: Better Regulation*

### 3.2. Evaluation principles

**Evaluation principles guide the entire evaluation policy, starting from institutional arrangements, as well as the conduct of internal or external evaluators, contracting authorities, and line managers**.

Four sets of general evaluation principles can be referred to: ethical principles, the general principles stated by the Better Regulation guidelines[14], the OECD DAC principles, and the principles which can be extracted from ENI[15] and IPA II[16] regulations. The present guidelines combine these principles.

The full list of relevant evaluation principles is presented in the **Box 3: Evaluation Principles**. Evaluation principles may refer to *ethics* (like respect for autonomy) on the one hand or to *methodological* or *practical* issues (like independence, openness, responsibility) on the other hand.

**Box 3: Evaluation Principles**

***Ethical principles***

**Respect for autonomy**: "autonomy is the 'right to act freely, make free choices, and think as you wish'.

**No maleficence**: The "do no harm" principle has a *negative* part (abstain from inflicting discomfort or injuries to others, e.g., when conducting interviews with refugees, thread carefully about sensitive issues) and a *positive* part (protect others from the risk of harm, e.g., by not disclosing respondents' identity when reporting on survey results).

**Beneficence**: help others, e.g., by providing information to help programme managers to make decisions.

The **Justice** has two dimensions: *procedural* justice (fairness in making decisions) and *distributive* justice (fairness in outcome allocations).

**Fidelity**: "keeping promises, being loyal, and being honest".

***Methodological and practice principles***

**Independence and Objectivity** (Better Regulation, OECD DAC Principles*)*. An evaluation can be considered independent when evaluators: (i) carry out their tasks without influence or pressure from the organisation launching the evaluation or being evaluated; (ii) are given full access to all relevant information; and (iii) have full autonomy in conducting and reporting their findings. Independence results from institutional arrangements that allow both **freedom to** conduct the research and express judgments and **freedom from** interferences and pressures. **Outsourcing evaluations to external consultants does not ensure independence. Appropriate institutional arrangements effectively shelter internal evaluators**. Four conditions must be fulfilled:

- Organisational independence: the evaluation unit should respond directly to executive boards (in organisations) or elective assemblies, rather than to a line unit. The internal and external evaluators should be functionally independent from the authorities responsible for intervention implementation.
- behavioural independence attains to individual attitudes and ethics as well as to the freedom of an evaluation unit to self-select the work programme and to manage its budget;
- avoidance of conflicts of interest; and
- Protection from external influence: ability to develop judgments freely, without fearing retribution.

**Openness** refers to knowledge sharing and to democratic accountability. EU regulations establish this principle, by requiring that all evaluations be public and that evaluations must be brought to the attention of the European Parliament, the Council, and the Member States.

[14] Ibidem.

[15] REGULATION (EU) No 232/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 11 March 2014 establishing a European Neighbourhood Instrument.

[16] REGULATION (EU) No 231/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 11 March 2014 establishing an Instrument for Pre-accession Assistance (IPA II).

**Responsibility** implies both the ability to recognise organisational, professional, or technical limitations (e.g., by requesting additional expertise when needed or by appropriately using an internal evaluation) and the ability to express one's positions in controversies.

**Comprehensiveness** (Better Regulation): The definition of evaluation targets five criteria – effectiveness, efficiency, relevance, coherence and EU added value. Other criteria may also be added as appropriate.

**Proportionality** (Better Regulation): The scope and analysis conducted must be tailored to the particular intervention, its maturity and the data available. For some criteria new data will need to be collected, analysed and compared with other findings; for others, a short summary can be presented based on existing reports and information or providing a standard explanation (e.g. in areas where the EU has exclusive competence).

**Credibility** (OECD-DAC Principles) depends on the transparency of the evaluation process and on evaluators' expertise and autonomy. Recipient countries' full participation in evaluation promotes credibility and commitment.

**Transparent Judgement** (Better Regulation): Evaluators must make judgements based on all evidence (good or bad). Judgement criteria for each evaluation question (success factors, related indicators, required evidence and information) should be clearly identified during the evaluation design. **Transparency** of the evaluation process is crucial to its credibility and legitimacy (OECD DAC).

**Evidence-based** (Better Regulation): Evaluations are based on the best available evidence (factual, opinion based etc.), drawn from a diverse and appropriate range of methods and sources (triangulation). Not all sources of evidence are equally robust and consideration must be given as to when and how the evidence was collected and whether there is any bias or uncertainty in it. Where possible, sensitivity and/or scenario analysis should be conducted to help test robustness of the analysis. Any limitations to the evidence used and the methodology applied, particularly in terms of their ability to support the conclusions, must be clearly explained.

**Usefulness** (OECD-DAC Principles): To have an impact on decision-making, evaluation findings must be perceived as relevant and useful, be presented in a clear and concise way, fully reflect the interests and needs of the many parties involved, be timely and easily accessible.

**Participation of donors and recipient** (OECD-DAC Principles): donors and recipients have to be involved in the evaluation process. Terms of reference should address issues of concern to each partner. Impartiality and independence equally apply to recipients and donors. Affected groups' views and expertise form an integral part of the evaluation. This principle is linked to:

- **Ownership** principle of ENI and IPA II beneficiaries.
- **Stakeholders' involvement**: affected parties must be involved in evaluation and **joint evaluations** should be undertaken.

**Donor co-operation** (OECD-DAC Principles): Joint donor evaluations should be promoted. This serves two goals: improving mutual understanding and reducing administrative burdens on recipients. Donors should exchange evaluation plans systematically and well ahead of actual implementation.

*Source: DG NEAR, adapted from* Better Regulation and OECD DAC Principles.

## 5. EVALUATION CRITERIA

Evaluation criteria are expressed as a relationship between one key element of an intervention (achievements or objectives) and another aspect, which, in turn, can be internal (e.g. resources, or activities) or external (e.g., needs or durability).

For example, objectives are judged on the basis of how well they reflect societal needs (*relevance*), while achievements on how well they correspond to objectives (*effectiveness*) or to societal needs (*utility*) or to a concept of equity (*equity*). Of course, the application of criteria is not automatic: evaluators will need to specify and explain their judgment.

There are three converging sources for criteria for DG NEAR evaluations:

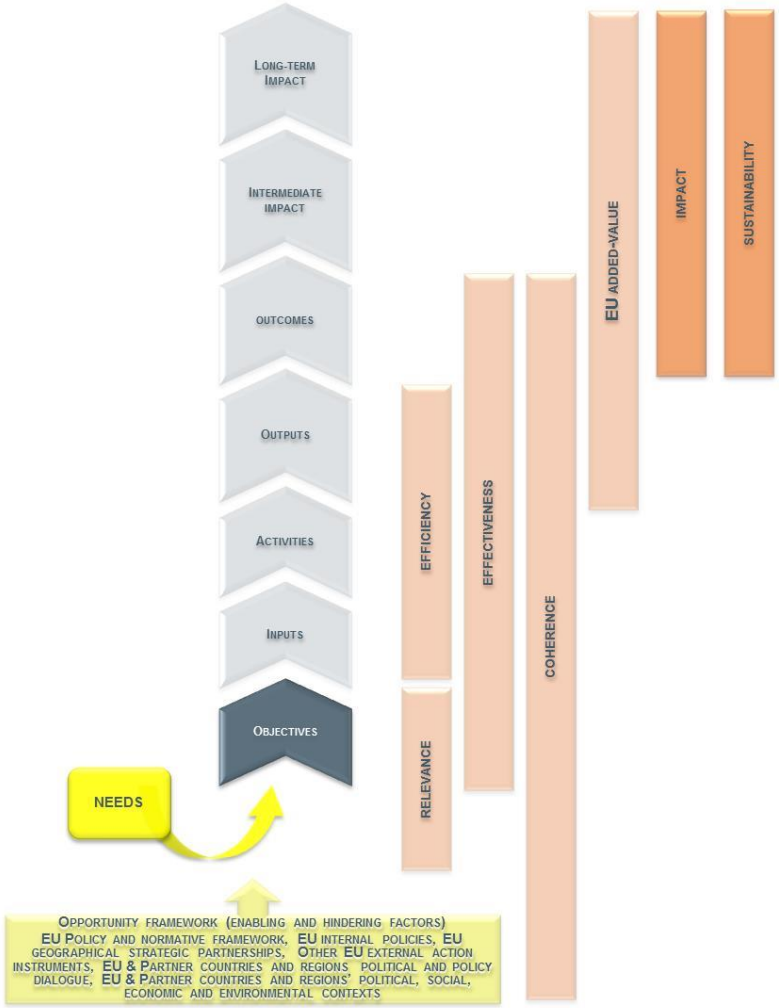
a) the Regulations governing external action in the framework of ENI[17] and IPA II[18]

b) the Better Regulation Guidelines[19]

c) The OECD-DAC criteria.

Three of the criteria (*Relevance, Effectiveness, and Efficiency*) are common to all sources. *Impact* is mentioned as a criterion among OECD DAC criteria and mandated by Regulations. *Sustainability* is explicitly mentioned in Better Regulation and OECD DAC, whereas *Coherence*, *EU-added value*, and *Equity* are mentioned both in Better Regulation and in Regulations.

**FIGURE 2: THE INTERVENTION LOGIC AND THE KEY 7 (5 BEING MANDATORY) EVALUATION CRITERIA**




*Source: DG NEAR; adapted from the Better Regulation*

The Better Regulation Guidelines sets as mandatory the following five criteria that have to be addressed in each evaluation: Relevance, Efficiency, Effectiveness, Coherence and EU added value (refer to **BOX 4: MANDATORY EVALUATION CRITERIA AS PER THE BETTER REGULATION GUIDELINES, WITH TYPICAL EXAMPLES OF QUESTIONS**). If these

[17] REGULATION (EU) No 236/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 11 Mar 2014 laying down common rules and procedures for the implementation of the Union's instruments for financing external action.

[18] COMMISSION IMPLEMENTING REGULATION (EU) No 447/2014 of 2 May 2014 on the specific rules for implementing Regulation (EU) No 231/2014 of the European Parliament and of the Council establishing an Instrument for Pre-accession assistance (IPA II).

[19] http://ec.europa.eu/smart-regulation/guidelines/toc_guide_en.htm.

are not covered, due justification must be provided. Additional criteria beyond these five can also be added (refer to **BOX 5: OTHER EVALUATION CRITERIA, WITH EXAMPLES OF TYPICAL QUESTIONS**).

It must be highlighted that the Better Regulation does not fully address the specificities of EU external action instruments. As such, EC HQs/EUDs/national authorities need to have an open mind in relation to other relevant criteria outside this mandatory list. This is especially the case of the remaining two OECD-DAC criteria: sustainability and impact (refer to **BOX 5: OTHER EVALUATION CRITERIA, WITH EXAMPLES OF TYPICAL QUESTIONS**).

**BOX 4: MANDATORY EVALUATION CRITERIA AS PER THE BETTER REGULATION GUIDELINES, WITH TYPICAL EXAMPLES OF QUESTIONS**

***Relevance*** looks at the relationship between the needs and problems in society and the objectives of the intervention. Things change over time - certain objectives may be met or superseded; needs and problems change, new ones arise. The OECD DAC attributes a wider meaning to this criterion, including also the correspondence between instruments and goals—a set of questions which is fundamental in any evaluation and which may overlap with the criterion of internal coherence.

Typical examples of relevance questions*:*

- *To what extent have the (original) objectives proven to be appropriate for the intervention in question?*
- *How well do the (original) objectives (still) correspond to the needs and priorities of the target groups (incl. governments)?*
- *How well adapted is the intervention to subsequent changes in the context in which the intervention is framed (changes in needs, in policies, etc.)?*
- *To what extent is the intervention aligned with the policies and strategies of the partner country?*

***Efficiency*** considers the relationship between the resources used by an intervention and the changes it generates (which may be positive or negative). Resources include staff, purchases, time and money spent, fixed costs, running costs, administrative and regulatory burden[20].

Differences in the way an intervention is approached and conducted can have a significant influence on the effects, making it interesting to consider whether other choices achieved the same benefits at less cost (or greater benefits at the same cost).

Typical examples of efficiency questions*:*

- *To what extent are the costs involved justified, given the changes/effects which have been achieved?*
- *To what extent are the costs proportionate to the benefits achieved? What factors are influencing any particular discrepancies?*
- *Which are the other ways of using resources that have produced more results or have used resources sparingly, yet maintaining the same level of achievements?*

***Effectiveness*** relates to the extent to which the intervention's objectives were achieved, or are expected to be achieved (OECD-DAC).

Typical examples of effectiveness questions*:*

- *To what extent have the objectives been achieved?*
- *What have been the (quantitative and qualitative) effects of the intervention?*
- *To what extent do the observed effects correspond to the objectives?*
- *To what extent did different factors influence the achievements observed?*

***Coherence*** is associated with two meanings: i) Internal coherence, looking at how the various internal components of an intervention operate together to achieve its objectives; and ii) External coherence, looking at how interventions within the same area and/or policy work together. At its widest, external coherence can look at compliance with international agreements/declarations.

Typical examples of coherence questions*:*

*Internal:*

[20] The Better Regulation toolbox includes a section addressing the issue of assessing the regulatory burden, European Commission, 2015b: 278.

- *How do the intervention components fit in with one another and with the objectives?*

*External:*

- *To what extent is this intervention coherent with other interventions that have similar objectives?*
- *To what extent is the intervention coherent with international obligations?*
- *To what extent is the intervention coherent with wider EU policy?*

***EU added value.*** Under the principle of subsidiarity (Article 5 Treaty on European Union), the EU should only act when the objectives can be better achieved by Union action rather than by potentially varying action by Member States. It requires consideration of the value and improvements which are caused by the EU rather than another party taking action.

Typical examples of EU added value questions

- *Is the EU support generating better results than what would happen without it?*
- *What is the additional value resulting from the EU intervention(s), compared to what could be achieved by others (i.e. EU Member States, other donors, and the country/region)?*
- *To what extent do the issues addressed by the intervention continue to require EU support?*
- *What would be the most likely consequences of stopping or withdrawing the existing EU intervention?*

*Source: DG NEAR, adapted from* Better Regulation and OECD DAC Principles

As recalled by the Better Regulation, there are also several further evaluation criteria that it may be appropriate to consider, depending on the type of intervention and the timing of the evaluation. The most common additional criteria used by the Commission are shown in **Box 5: Other Evaluation Criteria, with examples of typical questions.**

**Box 5: Other Evaluation Criteria, with examples of typical questions**

***Utility*** relates to the extent to which an intervention fulfils societal needs. Questions inspired by this criterion put together findings from analyses under the criteria of **relevance** and **effectiveness**, which are high in priority in all the sources relevant for the work of DG NEAR. Utility questions are particularly interesting when dealing with interventions which are repeated in a similar way from country to country or which have been in operation in the same area for longer than an intervention period.

Typical examples of utility questions*:*

- *To what extent do the changes/effects of an intervention satisfy (or not) stakeholders' needs?*
- *How much does the degree of satisfaction differ according to the different stakeholder groups?*

**Impact** relates to the changes that are expected to happen due to the implementation of an intervention. Such impacts may occur over different timescales and affect different actors. They can be positive and negative, direct and indirect, intended or unintended, on any dimension (social, economic, environmental, political, etc.).

Typical examples of impact questions*:*

- *Which changes has the intervention produced, regardless of its objectives?*
- *Which of these changes are long-term?*
- *Which changes have affected which groups?*

**Sustainability** relates to the continuation of benefits from an intervention after major support has been completed. The probability of continued long-term benefits. The resilience to risk of the net benefit flows over time (adapted from the OECD-DAC). It has various dimensions: social, economic, political, environmental, financial, institutional, etc.

Typical examples of sustainability questions*:*

- *How likely the effects are to last after the intervention ends?*
- *Which institutional arrangements allow for maintaining the benefits achieved?*
- *Which financial resources are available to fund the continuation of the services provided by the intervention? How long are they likely to be available and from which sources?*

- *How is the scale of relevant political entities aligned with the level of the consequences of their decisions?*
- *How has the environment been modified by the intervention?*
- *Which are the environmental consequences of delivery of services made possible by the intervention?*

**Equity** relates to the capacity of an intervention to redress societal, gender, or territorial inequalities—in other words, equity refers to the so-called **crosscutting issues** that have to be taken into consideration in all EU policies.

Typical examples of equity questions*:*

- *How do intervention activities redress situations in which a group suffers discrimination?*
- *How fairly are the different effects distributed across the different stakeholders / regions? / Genders? / Social groups?*

**Complementarity**. The Better Regulation guidelines define this criterion as the connection between EU interventions and Member States' policies, while referring to internal EU policies. With regard to DG NEAR, this criterion should be interpreted to mean complementarity with i) other donors' initiatives in the partner country(ies) and with ii) the partner national policies.

Typical examples of complementarity questions*:*

- *Which are the EU policies and interventions' interactions with policies that Member States and other development partners implement in the partner countries?*
- *To what extent do EU policies and interventions support and usefully supplement other policies (in particular those pursued by the Member States)?*
- *To what extent was the EU-supported intervention aligned with Government priorities? How did it respond to changing national priorities over the evaluation period?*

**Coordination** relates to the degree of organisation and link among interventions in order to take into account their interactions

Typical examples of coordination questions*:*

- *To what extent are interventions in the same sector or area organised to maximise their joint effects, e.g. by mobilising resources combined with harmonising measures?*
- *Which organisational arrangements are in place in order to maximise the combined effect of different interventions in the same sector or area?*

*Coordination can also be internal, within a single intervention.*

A typical example of internal coordination question is*:*

- *Which institutions are included in the process/intervention and how frequent is the mutual communication between the relevant stakeholders?*

**Acceptability** relates to the changes in stakeholders' and public's perception of an intervention and, therefore, to the type of support it engenders or resistance it may produce.

Typical examples of acceptability questions*:*

- *To what extent can we observe changes in the perception of the intervention (positive or negative) by the targeted stakeholders and/or by the general public?*
- *Which type of support has the intervention engendered? Which opposition? By which groups?*

*Source: DG NEAR, adapted from* Better Regulation and OECD DAC Principles.

## 6. REASONS AND USES OF M&E

### 6.1. Why monitor? Why evaluate?

**Monitoring** is a management tool. Monitoring gives regular and systemic information on *where* an intervention is at any given time (and over time) relative to respective targets and outcomes (and related

indicators). It is descriptive in intent[21]. It aims to identify successes, problems and/or potential risks so that corrective measures are adopted. Even though it focuses mainly on the intervention's inputs, activities and outputs, it should also look at how the outputs can effectively induce the outcomes that are aimed at[22].

**Evaluation** gives evidence of *why* intended changes are or are not being achieved. It seeks to address issues of causality[23]. It has multiple purposes.

- First, **learning from experience**. It provides organisations with actionable knowledge and lessons learnt they can use to improve their policies and actions: "an evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision–making process of both recipients and donors"[24]. This has two facets: on the one hand, evaluation provides elements to decide on policies and actions and a basis for resource allocation. On the other hand, it provides knowledge for improving implementation and the relative decision-making. The Better Regulation guidelines also recognise this function. [25 & 26]

  In this framework:

  - Evaluations can play a **persuasive role**: organisations can use the knowledge they acquire through evaluation in negotiations with other organisations and with partner country counterparts at various levels and to demonstrate results. Within a same organisation, programme managers can use evaluation findings in internal negotiations. For example within the Commission and with Headquarters, arguing for the right resources and appropriate timing. They can use evidence from evaluations in dealing with partner country counterparts.
  - Evaluation produces **knowledge development**. This function interests planners and policy-makers, policy communities, and citizens.
- Another fundamental function of evaluation is **accountability**. Within the European Commission, the primary functions of evaluation are to support the Commission work and to provide information to the European Parliament and to stakeholders, at all levels: at the EU level (Council, Court of Auditors, etc.), at Member State level (national parliaments, taxpayers, NGOs, etc.) and at the partner country level (governments, national parliaments, citizens, social and economic interest groups, civil society organisations, etc.).
- Programme managers can use evaluation findings in order to demonstrate results, to show evidence of the difficulties they encounter, and to provide proof of the solutions they implemented.

[21] Jody Zall Kusek & Ray C. Rist, Ten Steps to a Results-Based Monitoring and Evaluation System – A Handbook for development practitioners, Global HIV/AIDS Program and Operations Evaluation Department, the World Bank, 2004, p. 13.

[22] http://ec.europa.eu/europeaid/monitoring-projects-and-programmes_en

[23] Jody Zall Kusek & Ray C. Rist, 2014: 13.

[24] OECD DAC, 2010a: 21.

[25] "More often, several evaluations or evaluations combined with other evidence and opinion are used cumulatively to inform debates and influence decision-making. Evaluation thus stimulates the process of debate, challenge and counter challenge to evidence and its interpretation" (European Commission, 2013: 57).

[26] "Evaluation is not the end of the process. Completing the final report and disseminating findings should stimulate discussion of the evaluation findings. In turn, this should lead to the identification of appropriate follow-up actions to put into practice the lessons learned and feed the evaluation findings into the next cycle of decision making" (European Commission, 2015c: 247).

Also, it is important for the focal points in national authorities to secure follow up of the achievements of evaluators' recommendations towards the beneficiaries/target groups, in order to assist them in their process of acquiring the needed evaluation standards. In the case of IPA, the results of such evaluation findings and recommendations should be recorded at the IPA Monitoring Committees.

- In this framework evaluations may produce **institutional strengthening**, by allowing public organisations to communicate better their results and be more open to inputs from beneficiaries and civil society.
- As mentioned under Consultation strategy, **evaluations must be informed and informing exercises**. By allowing key stakeholders to be part of the entire evaluation process (from design to implementation, depending on the type of stakeholder), evaluations are to be considered as key paths towards **ownership**, therefore contributing to sustainability.
- Finally, there is a **compliance** function. Evaluation is an opportunity. This is often obscured by the fact that it is a mandated, potentially threatening activity, undertaken to fulfil legal requirements. This function may help develop evaluation practices that grow to play the other functions as well. However, it may also result in token activities whose results are little used.

These evaluation uses, mainly the first three, are interconnected. And if applied to the entire portfolio of evaluations commissioned by the EC, they would **contribute to the institutionalisation of the evaluation function**. As synthetized by the Independent Evaluation Group of the World Bank, the intervention logic of institutionalizing evaluation would have five levels: It starts with the value addition that can be generated when evaluations complement each other, create a deeper understanding, and lead to institutional learning and accountability. Or, an institution's credibility may receive a boost when it can demonstrate to stakeholders such as civil society and partners that it has a strong evaluation function and learns from its insights. Management benefits from independent evaluation in governing the institution with an eye to increase development effectiveness, which benefits partner countries and ultimately their citizens[27].

**FIGURE 3 : THE INTERVENTION LOGIC OF INSTITUTIONALIZING EVALUATION**


*Source: adapted from https://ieg.worldbankgroup.org/blog/institutionalizing-evaluation-what-theory-change*

Evaluation use of any kind, albeit desirable, does not happen automatically. Rather, **in order to ensure the use of evaluation findings, evaluation managers have to carefully arrange key moments of the evaluation process**, starting from the preparation phase. Most of the use of evaluation findings does not happen as a consequence of reading reports but, rather, as **process use**. The self-reflection induced by the necessity of launching and managing an evaluation and the interaction with evaluation managers and evaluators change programme managers' outlook on their intervention and on the results they are producing[28]. This, of course, does not reduce the importance of evaluation reports and of other ways of communicating evaluation results.

[27] Adapted from https://ieg.worldbankgroup.org/blog/institutionalizing-evaluation-what-theory-change

[28] "Even where evaluation results are not used the process of evaluation initiation and reflection can be useful by offering opportunities to exchange information and clarify thinking" (European Commission, 2013: 57).

### 6.2. Users

The intended users of evaluation are those individuals or groups who have a vested interest in the evaluation results and are in a position to make decisions or take action based on the evaluation results[29]. As such, they have a stake in the evaluation recommendations and/or responsibilities in implementing them. Users can be: policy makers and intervention designers, managers and operators in charge of the implementation, partners, institutions having provided financing and to whom accountability is required, public authorities conducting connected or similar interventions, civil society organisation, and experts[30].

By extension, this same definition can be applied to monitoring users. The only distinction being that users are mostly confined to the intervention's management level.

### 6.3. Connection between monitoring and evaluation M&E

#### *6.3.1. Two separate yet complementary areas of practice*

Evaluation and monitoring are two separate areas of practice, providing different inputs into the decision-making process at different points in time. They are, however, interlinked: monitoring benefits from evaluations and, in turn, evaluations use data coming from monitoring. Programme managers can use combinations of these two forms of organising and eliciting information about interventions to improve planning and implementation.

**TABLE 1: M&E, TWO SEPARATE YET COMPLEMENTARY AREAS OF PRACTICE**

| | Monitoring | Evaluation |
|---|---|---|
| **Objective** | To identify implementation problems<br>To assess the progress towards objectives<br>To check whether the M&E system is in line with the approved design | Learning and accountability<br>To assess actual outcomes and impacts |
| **Timing** | Continuous, all along the intervention's duration. It produces data in time series. | In specific moments of the intervention's cycle |
| **Users** | EC/EUDs and national/regional programme managers | EC/EUDs and national/regional authorities, other key stakeholders (i.e. beneficiaries) and the general public<br>The European Parliament and the Council are the ultimate audience of reports[31] |
| **Responsibilities** | Operational staff and/or external monitors | Internal and/or external evaluators |
| **Focus** | On implementation:<br>• Interventions' progresses and first results.<br>• Inputs, activity execution and contribution to outputs<br>• Efficiency and effectiveness<br>Key question: | On results:<br>• Outcomes and intermediate and long-term impacts.<br>• Strategic aspects (relevance, coherence, EU added value) and sustainability<br>• Explores unintended results<br>Key question: |

[29] United Nations Development Programme, 2009: 129.

[30] Methodological basis for evaluation, Joint Evaluation Unit, DG EuropeAid, EC, 2006.

[31] "The Commission shall send its evaluation reports to the European Parliament, to the Council and to the Member States through the relevant committee referred to in Article 16. Specific evaluations may be discussed in that committee at the request of Member States. The results shall feed back into programme design and resource allocation" (Regulation (EU) No 236/2014 of the European Parliament and of the Council, 11/03/2014, Art. 12).

| | Monitoring | Evaluation |
|---|---|---|
| | **What** is the intervention doing and how much implementation is progressing? | **To what extent** and **why** is the intervention producing (or failing to produce) the specific outcomes and impacts which have actually materialised—be those negative or positive? |
| **Methods** | Greater emphasis on quantitative indicators<br>Output and process indicators<br>Quick exercise gathering data from systematised internal systems[32] | Greater emphasis on qualitative indicators<br>Outcome and impact indicators<br>Rigorous and sophisticated exercise, gathering data emerging from external sources, such as research |
| **Reports** | Standard and comparable formats<br>Report progress | Less standardised presentation,<br>In-depth analysis on selected issues linked to the achievements of intended objectives<br>Provide lessons and recommendations |
| **Dissemination** | Normally limited to intervention's staff and direct users | More wider, including relevant donors' services and the general public |

*Source: DG NEAR, adapted from various sources*

### *6.3.2. When to use monitoring and when to plan evaluations*

Normal monitoring activities are compatible with whatever complementary control tool. What needs to be considered while planning an activity is basically that ROM and mid-term evaluations are alternative tools. When the two are considered the decision on whether to go for the first or the second depends on the actual needs (considering that the scope and the depth of analysis are different, as showed in TABLE 2: GUIDANCE TO THE SELECTION OF THE APPROPRIATE CONTROL TOOL) and to a lesser extent on cost/effectiveness considerations, a ROM being quicker and cheaper than an evaluation.

The following represent some criteria based on which it can be decided on the relevant control tool to use with regard to actions.

**TABLE 2: GUIDANCE TO THE SELECTION OF THE APPROPRIATE CONTROL TOOL BY DG NEAR**

| Criteria | Regular monitoring activities | On the spot checks | ROM | Mid-term evaluation | Audit |
|---|---|---|---|---|---|
| Low overall risk | ***** | * | * | * | * |
| High operational risk | ***** | **** | ***** | **** | * |
| High contractual/financial risk | **** | **** | ** | * | **** |
| Evaluation budget not available | *** | *** | **** | * | * |
| Lack of sector expertise at EUD | * | * | **** | * | * |
| High workload | * | * | **** | *** | ** |
| Highly innovative Action | **** | * | **** | *** | ** |
| Highly sensitive Action | ***** | * | ***** | *** | * |
| ROM mission already scheduled | ** | * | ***** | * | * |
| Mid-term evaluation already scheduled | ** | * | * | ***** | * |
| Action just started | ** | * | * | * | ** |
| Action in the final phase - no chance of extension | *** | * | * | * | ** |

[32] Monitoring systems may also include data coming from external sources, such as surveys. Particular care should be used when providing for ad hoc surveys to feed a monitoring system, since they require great effort and expenses and risk being performed more rarely than foreseen. It is usually advisable to build monitoring systems around administrative data and, for other data, to rely on data that are periodically produced by specialised organisations (e.g., a country's statistical office).

| Criteria | Regular monitoring activities | On the spot checks | ROM | Mid-term evaluation | Audit |
|---|---|---|---|---|---|
| Action experimenting operational problems: necessity of a quick fix | *** | *** | ***** | * | * |
| Discussion on possible extension of an Action close to the end[33] | *** | ** | *** | * | * |
| Necessity to decide on possible extension based on solid evidence: time still available | *** | * | ** | *** | * |

Legend: adequacy is defined by the number of asterisk: 1 low; 5 very high.

*Source: DG NEAR*

Whatever control tool is selected, it is worth mentioning that NEAR MIS (Management Information System), to be used by EC HQ's and EUD staff during the entire Action's lifecycle, is also expected to provide a monitoring platform in which key performance indicators (a feature newly introduced) are expected to be reported on and used as sources of information for the different control tools.

### *6.3.3. From monitoring to evaluation*

**Evaluation uses data from monitoring**. A good monitoring helps minimise the need of *ad hoc* data collection. Evaluators may use data, monitoring reports and minutes of relevant monitoring platforms reporting on:

- **Inputs and expenditures** in analyses about efficiency;
- **Activities**, in order to study implementation and to analyse efficiency;
- **Outputs,** for analyses of what the intervention has produced. This is a first step in understanding contribution to outcomes;
- **Outcomes,** as a starting point in the determination of causal links or contribution.

Evaluations, however, cannot solely rely on monitoring data because:

- Monitoring data only concern the first levels of an intervention logic: outcome and impact remain substantially out of the reach of a monitoring system, even if contribution to outcomes can also be covered, depending on the timing in which monitoring was carried out.
- Monitoring data only detect what they were set up to study. Indeed, monitoring data emerges from specific indicators defined during planning/programming[34] and included in the M&E system. Unexpected twists in implementation risk therefore being undetected. In particular, the important data about which mechanisms operated between the different levels of an intervention logic are not covered in monitoring systems.

Due to these constraints, in any evaluation, **evaluators** need to complement data from monitoring and other administrative sources with data about mechanisms, outcomes and impacts, collected from reliable external sources and from original field research. **Evaluation managers** need to ensure that they do so, by explicitly requesting such research in Terms of Reference (ToR) and throughout the evaluation.

[33] The ROM Handbook does not foresee this case (indeed, it does not foresee a ROM mission in the last six months) but in exceptional circumstances this can be useful.

[34] Refer to § **8.3.2 STEPS: DEFINING THE INTERVENTION LOGIC'S DIAGRAM (STEPS 1 TO 5), DEVELOPING THE INTERVENTION LOGIC'S NARRATIVE (STEP 6)** & & for further details.

#### *6.3.4. From evaluations to monitoring*

First, **evaluation produces knowledge to build or improve monitoring systems**, based on the Better Regulation's "Evaluate first" principle.

Evaluations provide knowledge about the **linkages between the levels of an intervention logic**: causal chains, contribution connections, or permanent correlations. Programmers use this knowledge in identifying/revising indicators and in building/strengthening monitoring systems. Evaluations include this information even though they are not focused on the functioning of the monitoring system.

While monitoring says whether or not progress is on track, evaluations assess the reasons underpinning such progress (or lack of)[35].

Evaluations may provide knowledge on how monitoring arrangements have worked (i.e., who produced the data, who collected and entered it, who cleaned it, who used it and for what, and, finally, which conflicts arose around monitoring). This is important, since some of the mistakes may hide conflicts about monitoring (e.g., a low priority on monitoring leading to contracting it out or to staffing it with untrained workers) or programmes (e.g., contrasting interpretations on targeting). Evaluations provide this information, however, only if evaluation questions included a critical exam of M&E arrangements. Exploring the link between monitoring and evaluation is a very important step in an intervention's evaluation, in order for the evaluation to provide conclusions on the functioning of the monitoring system and findings that may be used to enhance its capacities. Therefore, it is recommended to include questions on the monitoring system in relevant evaluations.

## 7. MAIN GENERAL RESPONSIBILITIES FOR M & E

### 7.1. Monitoring actors

The following table presents who the different actors responsible for monitoring are in the framework of DG NEAR's mission.

**TABLE 3: MONITORING ACTORS AND RELEVANT KEY RESPONSIBILITIES**

| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| | **EC HQs & EUDs** | | |
| **DG NEAR management** | Both internal and external monitoring (ROM). Under the initiative of the M&E service:<br>• Ensures that monitoring principles are respected<br>• Determines what is to be monitored<br>• Allocates adequate financial and human resources for collection, analysis, reporting and use of monitoring results<br>• Ensures the up-take of monitoring findings by EC HQs/EUD staff | ✔ | ✔ |
| **DG NEAR M&E Service** | Both internal and external monitoring (ROM)<br>• Set up standards<br>• Ensures proper implementation of guidelines<br>• Promotes knowledge management and organisational learning<br>• Provides, upon request methodological guidance on:<br>  • Monitoring tools (on the spot checks checklists, etc.)<br>  • Reporting | ✔ | ✔ |
| **Monitoring focal point in EC HQ's Units & EUDs** | Under the coordination of the Head of Unit/of Operations, (s)he **coordinates** the following activities:<br>Both internal and external monitoring (ROM): | ✔ | ✔ |

[35] Outcome indicators detect changes in dimensions which are initially believed to be connected with interventions. It is, however, impossible to claim that there was a link between the intervention and the type or size of the outcome without an evaluation.

| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| | • Ensures that monitoring principles are respected<br>• Coordinates the preparation and updating of monitoring plans<br>• Coordinates the preparation and updating of risk assessment<br>• Up-take of monitoring results by EC HQs/EUD staff<br>• Collection of inputs for relevant monitoring meetings and other reporting exercises<br>• Organisation and coordination of relevant follow-up<br><u>External monitoring (ROM):</u><br>• Organisation of ROM missions, in consultation with the ROM coordinator in DG NEAR HQ's Monitoring & Evaluation service. This includes ensuring that updated and quality checked monitoring data is provided to internal and external evaluators promptly and in a usable form.<br>Furthermore, (s)he acts as a repository of monitoring practice, lessons learnt, etc. in its duty service. | | |
| **Programme manager**[36] | For its actions' portfolio, (s)he:<br><u>Both internal and external monitoring (ROM):</u><br>*On planning and management*:<br>• Prepares and updates:<br>– Monitoring plans<br>– Risk assessments<br>• Prepares, manages and controls contractual documents, and prepares forecasts on contract payments.<br>• Supports timely disbursement of EU resources, based on approved work plans and budgets and an assessment of action performance.<br>*On internal monitoring systems*:<br>• Assesses the quality of actions' internal monitoring systems and where required, plans/implements support to improve them. This includes, among other:<br>• Provision for periodical "data cleaning:" check for mistakes and look for duplications, systematic misunderstanding, or missing data<br>• Support data users in understanding them: how they are collected, what they mean, to which need they respond to. This is particularly important for external evaluators, who have little knowledge of the way the monitoring system is built and of its logic.<br>• Allows for decentralised use of monitoring data by organisations that are responsible for collecting and providing data. This can help them understand the relevance and meaning of the data and, therefore, improve data collection, and reduce mistrust in the monitoring system<br>• Provides monitoring data to internal and external evaluators promptly and in a usable form. Update the data on demand and whenever there is a new release of monitoring data<br>• Under centralised management, entertains relations with contractors, in coordination with the C&F section<br>*On actual monitoring:*<br>• Contributes, as appropriate, to regular reviews of action progress and updating of operational plans through regular contact with action implementers and other donors.<br>• Keep appropriate records of action progress, the results achieved and constraints encountered.<br>• Comply with relevant instructions from DG NEAR Management on action's monitoring.<br>*On reporting (reports & meetings)*:<br>• Assess the content and quality of monitoring reports from implementing partners and suggest corrective measures, as required, to support efficient and effective | ✔ | ✔ |

[36] A programme manager refers to an EC officer who has an operational responsibility for overseeing and supporting the effective formulation, implementation and/or monitoring of specific actions financed by the EU, Strengthening project internal monitoring How to enhance the role of EC task managers, DG EuropeAid, European Commission, June 2007.

| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| | implementation<br>• Provides inputs/comments, as appropriate, for relevant monitoring meetings and other reporting exercises<br><u>Internal monitoring:</u><br>• Carries out internal monitoring activities and missions according to plan<br>• Prepares and updates internal monitoring note/reports<br><u>External monitoring (ROM):</u><br>• Ensures and supports the correct implementation of ROM missions, including:<br>  • Facilitating monitors' access to stakeholders and officials in partner countries, and potential and actual beneficiaries of the actions<br>  • Providing feedback on monitoring reports. Ensures the final draft meets quality standards<br>• Collaborate with ROM teams and follow up, as appropriate, on their recommendations | | |
| **Contract & finance staff** | • Accompany programme managers in the implementation of 'on the spot-checks'<br>• Accompany programme managers in the development of annual monitoring plans, etc. | ✔ | ✔ |
| **Head of Unit/Operations** | In coordination with EC DG HQ's Director, country/regional desk and Heads of Section, as well as with the national authorities, and with the support of the monitoring focal point:<br>• Ensures a smooth implementation of all internal and external monitoring activities<br>• Provides feedback to programme managers on monitoring findings and on relevant actions to be taken<br>• Ensures the actual up-take of monitoring results in the EC HQ/EUD work and on the implementing partner side.<br>• Provides for communication and dissemination of monitoring results, including best practices in terms of monitoring tools, use of monitoring results, monitoring processes, etc. | ✔ | ✔ |
| | • Ensures the preparation of the annotated agenda for IPA Monitoring and Sector Monitoring Committees | | ✔ |
| **Country/regional desk within Geographical Directorates** | • Supervises the implementation of financial cooperation and contributes to internal and external monitoring<br>• Reviews internal and external monitoring reports, as appropriate<br>• Provides inputs/comments, as appropriate, for relevant monitoring meetings and other reporting exercises. | ✔ | ✔ |
| | • Contributes to the preparation of the annotated agenda for relevant monitoring meetings at national level | | ✔ |
| **Other DGs** | • Contribute to internal and external monitoring, as appropriate<br>• Review internal and external monitoring reports, as appropriate<br>• Provide inputs/comments, as appropriate, for relevant meetings and other reporting exercises. | ✔ | ✔ |
| | **National authorities** | | |
| **Programme Management Units** | • Respect the relevant principles set up in **BOX 2: MONITORING PRINCIPLES AS DEFINED BY THE BETTER REGULATION**<br>• Review implementation of activities by means of performance indicators, including those set out in the Action Document, and other management tools<br>• Implement monitoring missions and provide feedback<br>• Provides monitoring data to internal and external evaluators promptly and in a usable form. Update the data on demand and whenever there is a new release of monitoring data<br>• Provides for periodical "data cleaning:" check for mistakes and look for duplications, systematic misunderstanding, or missing data<br>• Support data users in understanding them: how they are collected, what they mean, to which need they respond to. This is particularly important for external evaluators, who have little knowledge of the way the monitoring system is built and of its logic. | ✔ (as appropriate, lower intensity) | ✔ (as appropriate, lower intensity) |

| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| | • Organise Steering Committee meetings and prepare relevant reports<br>• Provide data on action implementation to the EUD/national aid coordinator, ROM monitors informing the agreed indicators, whenever relevant | | |
| **National aid coordinator** | • Oversees implementation of EU support actions<br>• Under decentralised management, coordinates monitoring activities of the different national institutions. This includes ensuring that updated and quality checked monitoring data is provided to internal and external evaluators promptly and in a usable form.<br>• Allows for decentralised use of monitoring data by organisations that are responsible for collecting and providing data. This can help them understand the relevance and meaning of the data and, therefore, improve data collection, and reduce mistrust in the monitoring system<br>• Ensures the actual up-take and follow-up of monitoring results at the national level | ✔ (as appropriate, lower intensity) | ✔ |
| **Contracting authorities** | • Supervise of EU support actions by means of contacts with the contractors and on the spot checks | ✔ | ✔ |
| **Line Ministries** | • Supervise implementation of EU support actions, as appropriate, also by means of monitoring missions<br>• Contribute to internal and external monitoring, as appropriate<br>• Review internal and external monitoring reports, as appropriate<br>• Provide inputs/comments, as appropriate, for relevant meetings and other reporting exercises.<br>• Uptake (use of) the monitoring findings<br>• Ensure the relevant follow-up | ✔ | ✔ |
| **External actors** | | | |
| **Other implementing partners (other than EC HQ/EUD staff and national authorities)** | • Respect the relevant principles set up in **Box 2: Monitoring principles as defined by the Better regulation**<br>• Establish actions' internal monitoring systems and provide for their update<br>• Provides for periodical "data cleaning:" check for mistakes and look for duplications, systematic misunderstanding, or missing data<br>• Support data users in understanding them: how they are collected, what they mean, to which need they respond to. This is particularly important for external evaluators, who have little knowledge of the way the monitoring system is built and of its logic.<br>• Provide monitoring data to internal and external evaluators promptly and in a usable form. Update the data on demand and whenever there is a new release of monitoring data<br>• Interact with EC HQs/EUD programme managers<br>• Carry out, in liaison with programme managers, internal monitoring missions<br>• Draft, present, and discuss internal monitoring note/reports<br>• Organise Steering Committee meetings and prepare relevant follow-up reports | ✔ | ✔ |
| **External monitors (ROM)** | • Respect the relevant principles set up in **Box 2: Monitoring principles as defined by the Better regulation**<br>• Interact with DG NEAR HQ's Monitoring & Evaluation service, and with programme managers<br>• Carry out, in liaison with DG NEAR HQ's Monitoring & Evaluation service and the monitoring focal point, ROM missions<br>• Draft, present, and discuss monitoring reports<br>• If needed, discuss with DG NEAR HQ's Monitoring & Evaluation service all instances of resistance or of interference. | ✔ | ✔ |
| **ROM contractor** | • Provide professional monitors/evaluators for whom previous work experiences are carefully checked<br>• Manage and supervise the monitor's work. As such, they are the final responsible actor on the monitor's work towards the service having launched the monitoring (EC HQ/EUD). In this framework, they are expected to:<br>  • Provide training/clear work instructions to the monitors before the monitoring starts. This should provide guidance to ensure that the monitor has | ✔ | ✔ |

| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| | a clear understanding of the tasks, of the process, the content and implications of the different steps.<br>• Allocate sufficient human and financial resources (out of the monitoring budget) to quality control (covering the quality of the process, of the monitors and of the different deliverables) | | |

*Source: DG NEAR and DG DEVCO (for some Programme managers' tasks)*

### 7.2. Evaluation Actors

Evaluation engages the same actors who are engaged in policy making and implementation. It, actually, provides the opportunity to also include actors who are key: potential and actual beneficiaries and intended target populations. All actors (especially evaluators, evaluation managers, contracting authority's representatives, programme managers) are bound by the ethical principles of evaluation[37].

The following table presents the main actors in evaluation and their respective roles in the framework of DG NEAR's mission.

**TABLE 4: EVALUATION ACTORS AND RELEVANT KEY RESPONSIBILITIES**

| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| | **EC HQs & EUDs** | | |
| **DG NEAR management** | <u>Both internal and external evaluations</u><br>• Ensures that relevant evaluation principles are respected<br>• Determines what is to be evaluated<br>• Allocates adequate financial and human resources for collection, analysis, reporting and use of evaluation results<br>• Provide clear information to the evaluation manager on how the findings will be used<br>• Defines the tasks, responsibilities, organisation and procedures for all actors involved in planning, designing and conducting evaluations, and disseminating and using evaluation results<br>• Linked to the latter, allocates adequate financial and human resources<br>• Promotes the use of evaluation in decision-making by ensuring that policy implications and lessons learnt from (and across) evaluations are synthesised and appropriately disseminated at EC HQs and EUD levels<br>• Ensures the relevant follow-up<br><u>External evaluations</u><br>• Protects the evaluators' autonomy and independence[38] | ✔ | |
| **DG NEAR M&E Service** | <u>Both internal and external evaluations</u><br>• Set up standards<br>• Ensures proper implementation of guidelines<br>• Prepares and implements strategic evaluation plan<br>• Coordinates preparation of operational evaluation plan by EUD and central operational units<br>• Ensures dissemination and publication of evaluation results; Promotes knowledge management and organisational learning<br>• Provides, upon request:<br>  • Methodological guidance on:<br>  • ToR elaboration, including on evaluators' selection criteria<br>  • Evaluation design | ✔ | ✔ |

[37] REFER TO 0 EVALUATION PRINCIPLES.

[38] Evaluators must have access to appropriate remedial actions and powerful actors in the case they encounter unjustified difficulties in performing the work in full autonomy (e.g., there are demands to limit the scope of field research or attempts to orient the choice of respondents) or are the object of undue pressures at any time during the evaluation.

| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| | • Quality assurance of the evaluation process and deliverables | | |
| **Evaluation focal point in EC HQ's Units & EUDs** | Under the coordination of the Head of Operations/ Unit::<br>• (s)he acts as intermediary between DG NEAR HQ's Monitoring & Evaluation service and programme managers in the field. In this framework (s)he:<br>  • Provides an overview of evaluation exercises in its duty service<br>  • Acts as a repository of evaluation practice, lessons learnt, etc. in its duty service<br>  • Collects requests on specific evaluation topics from programme managers<br>  • Advices programme managers, with the support of DG NEAR HQ's Monitoring & Evaluation service, on methodological aspects of each evaluation phase<br>• (s)he coordinates the following activities:<br>  • Preparation and updating of evaluation plans<br>  • Up-take of evaluation results by EC HQs/EUD staff<br>  • Collection of inputs for relevant meetings and reporting exercises<br>  • Organisation and coordination of relevant follow-up<br>• Furthermore, (s)he ensures that evaluation principles are respected | ✔ | ✔ |
| **Inter-Service Steering Group/Reference group** | Composed by representatives of EC HQs/EUDs (plus others DGs, for strategic evaluations), partner country/region authorities. It<br>• Steers the evaluation exercise in all key phases<br>• Provides input and information to the evaluation team and demonstrates an open and transparent approach to critically analysing performance and delivery<br>• Ensures the respect of principles set up in **Box 3: Evaluation Principles**. In this framework: the independence of evaluators and the credibility, impartiality and usefulness[39] of both the evaluation process and the final deliverable(s), are particularly important.<br>• Provide quality control on the different deliverables<br>• Interacts with evaluators on preliminary and final findings, taking care of avoiding to inadvertently impose interpretations, opinions, and values<br>• Reviews the different deliverables; ensures the final draft report meets quality standards<br>• Elaborate the quality assessment grid (as per the Better Regulation)<br>• Ensure a proper follow-up action plan after the completion of the evaluation | ✔ | ✔ |
| **Evaluation manager** | <u>Both internal and external evaluations</u><br>• Ensure the respect of principles set up in **Box 3: Evaluation Principles**<br>• Steers (provides operational support and coordination) the Inter-service consultative Group/Reference group.<br>• Call for periodic meetings to review progress, to identify solutions, and to discuss preliminary findings, on the basis of the draft final Reports and other deliverables[40]. In all occasions, ask questions and encourage action staff to ask questions.<br>• Request informal updates on research activities and preliminary results from evaluators, not limiting exchanges to mandated formal occasions.<br>• Help the evaluator understand the circumstances for each choice and action staff's standpoint.<br>• Debate with evaluators the obstacles they encounter in using the methods and techniques, with the goals of providing support to overcome them, of finding out about limits in administrative data availability and quality, and as an information source to build better knowledge on the limits and requirements of each method and technique.<br>• Involve, whenever possible, external experts and stakeholders in discussing | ✔ | ✔ |

[39] "It used to be common to regard the use of evaluation as being confined to acting on recommendations and final reports. It is now understood that evaluation use can be supported and occurs throughout an evaluation. Process use should involve stakeholders in evaluation thinking from the beginning. There are evaluations where the conclusions and recommendations are rejected but stakeholders, especially those involved in the steering committee find the evaluation useful. It can help them to clarify their own thinking and understanding and spark off innovative ideas for improvements. Promoting dialogue during the course of an evaluation is likely to ensure that when stakeholders receive reports they will be better prepared and receptive" (European Commission, 2013: 59).

[40] Participants in the mid-term meetings include at the minimum: the evaluator team, possibly including all team members, the evaluation manager, a representative of DG NEAR HQ's Monitoring & Evaluation service, when the evaluation manager doesn't come from that same service. The evaluator presents the preliminary findings and raises the most pressing issues. The evaluation manager leads the discussion. After the meeting, the evaluation manager is responsible for organising written feedback.

| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| | preliminary results, not to control the quality of the work of evaluators (this is the responsibility of the evaluation manager or of the Inter-service steering group/ Reference group), but, rather, to have different viewpoints on the substance of the action.<br>• Ensures the up-take of evaluation findings by EC HQs/EUD staff<br>• Collection of inputs for relevant meetings and reporting exercises<br>• Organisation and coordination of relevant follow-up, with senior level involvement (both at the level of EC HQs and EUD and partner country side)<br><u>Internal evaluations</u><br>• Prepares and updates internal evaluation reports<br><u>External evaluations</u><br>• Ensure the respect of principles set up in **BOX 3: EVALUATION PRINCIPLES.** In this framework: the independence of evaluators and the credibility, impartiality and usefulness of both the evaluation process and the final deliverable(s), are particularly important. If relevant, the evaluation manager should report any interference experienced by the evaluators to the relevant internal bodies within the Commission.<br>• Leads the development of the evaluation road map, when relevant[41], and provides for its publication<br>• Leads the development of the ToR<br>• Supervises the development of the consultation strategy (done by the evaluation teams), when relevant<br>• Manages the contractual arrangements, the budget<br>• Manages the selection and recruitment of the external evaluators<br>• Ensures and supports the correct implementation of external evaluations, including:<br>  • Evaluators' access to stakeholders and officials in EC HQs, EUDs and partner countries, and potential and actual beneficiaries of the actions. Ensures a fully inclusive and transparent approach to the evaluation<br>  • Accessibility of updated and reliable data (incl. that of monitoring) to external evaluators promptly and in a usable form. Involve programme staff in interactions with the evaluator and support them in providing data and materials and explaining the meaning of the data.<br>• Interacts with evaluators on preliminary and final findings, taking care of avoiding to inadvertently impose interpretations, opinions, and values<br>• Ask evaluators for detailed explanations of the methods and techniques they use and, as far as possible, involve programme staff in these meetings. This increases internal evaluation capacity and weakens resistance against evaluation findings.<br>• Support programme staff and partner country counterparts in interpreting evaluators' requests in a positive way, rather than defensively. For example, help them prepare for interviews by reviewing key material and reconstructing the history of the action. Help them interpret interviews as an opportunity to look at their work in a critical and analytical way.<br>• Reviews the different deliverables; ensures the final draft report meets quality standards<br>• Drafts the quality assessment grid of the final report, in consultation with the ISG/RG.<br>• Steers the elaboration of the Staff Working Document[42], when relevant, or the action plan<br>• Provides for follow-up and dissemination. | | |
| **Programme manager** | For its actions' portfolio, (s)he:<br><u>Both internal and external evaluations</u><br>• Prepares and updates evaluation plans<br>• Contributes to the development of the ToR, by identifying the purpose and scope of the evaluation<br>• Participates in the Inter-service consultative Group/Reference group | ✔ | ✔ |

[41] Project/programme evaluations managed by Delegations, or thematic evaluations launched at local level, do not need one.

[42] Refer to § **13.3.1.2 EVALUATION STAFF WORKING DOCUMENT.**

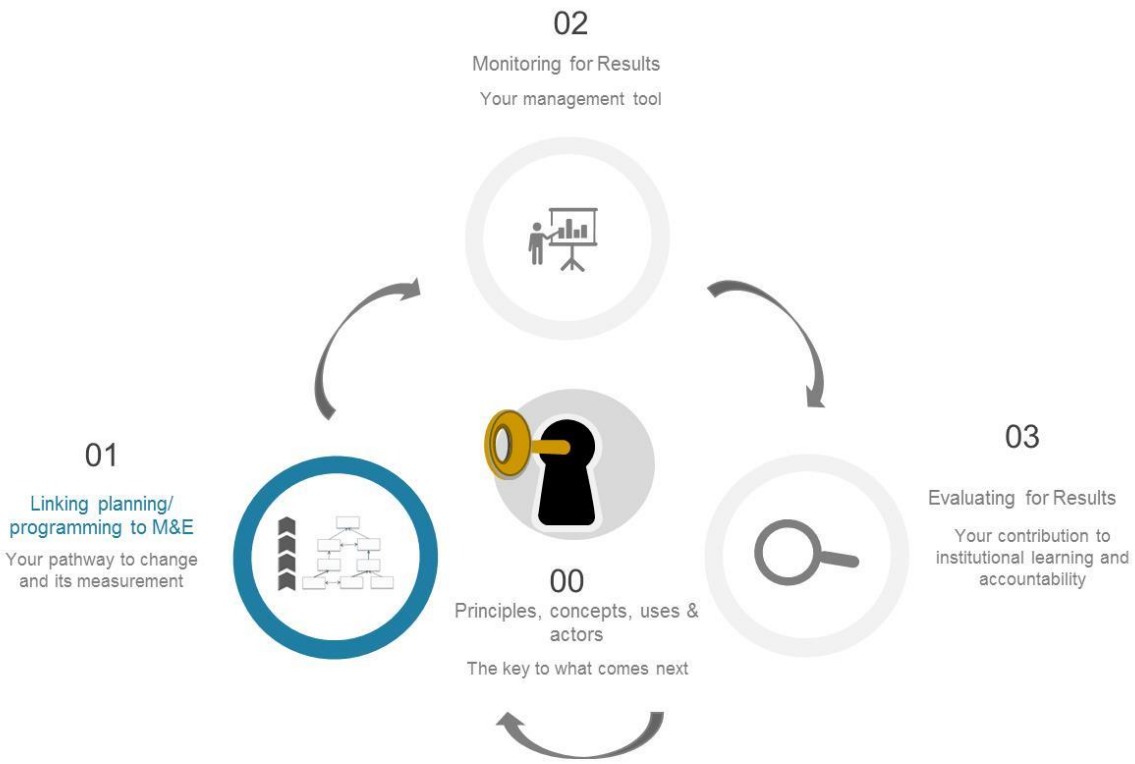
| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| | • Ensures the up-take of evaluation findings within its action and in the case of generalizable findings, informs relevant colleagues<br>• Collection of inputs for relevant meetings and reporting exercises<br>• Organisation and coordination of relevant follow-up<br>External evaluations<br>• Contributes to the development of the evaluation road map, as appropriate<br>• Ensures and supports the correct implementation of external evaluations, including:<br>  • Facilitating evaluators' access to stakeholders and officials in partner countries, and potential and actual beneficiaries of the actions<br>  • Providing updated and reliable data (incl. that of monitoring) to external evaluators promptly and in a usable form.<br>  • Providing feedback on inception report and draft and final evaluation reports.<br>• Support data users in understanding them: how they are collected, what they mean, to which need they respond to. This is particularly important for external evaluators, who have little knowledge of the way the monitoring system is built and of its logic.<br>• Allows for decentralised use of monitoring data by organisations that are responsible for collecting and providing data. This can help them understand the relevance and meaning of the data and, therefore, improve data collection, and reduce mistrust in the monitoring system<br>• Under centralised management, entertains relations with contractors, in coordination with the C&F section<br>• Provides inputs/comments, as appropriate, for relevant meetings and reporting exercises. | | |
| **Head of Unit/Operations** | In coordination with EC DG HQ's Director, country/regional desk and Heads of Section, as well as with the national authorities, and with the support of the evaluation focal point:<br>Both internal and external evaluations<br>• Supports the evaluation manager also ensuring that sufficient time can be allocated to the task<br>• Contributes to the development of the ToR, by identifying the purpose and scope of the evaluation<br>• Ensures a smooth implementation of all internal and external evaluation activities<br>• Provides feedback to programme managers on evaluation findings and on relevant actions to be taken<br>• Ensures the actual up-take of evaluation findings in the EC HQ/EUD work and on the implementing partner side.<br>• Provides for communication and dissemination of evaluation findings<br>External evaluations<br>• Contributes to the development of the evaluation road map, as appropriate<br>• Protects the evaluators' autonomy and the credibility of the evaluation | ✔ | ✔ |
| **Country/regional desk within Geographical Directorates** | • Supervises the implementation of financial cooperation and contributes to internal and external evaluation<br>• Reviews internal and external evaluation reports, as appropriate<br>• Provides inputs/comments, as appropriate, for relevant meetings and reporting exercises. | ✔ | ✔ |
| **Other DGs** | • Contribute to internal and external evaluations, as appropriate<br>• Review internal and external evaluation reports, as appropriate<br>• Provide inputs/comments, as appropriate, for relevant meetings and other reporting exercises. | ✔ | ✔ |
| | **National authorities** | | |
| **Evaluation manager** | Under indirect management:<br>• Ensures the respect of principles set up in **Box 3: Evaluation Principles** In this framework: the independence of evaluators and the credibility, impartiality and usefulness of both the evaluation process and the final deliverable(s), are particularly important.<br>• Drafts the Evaluation Plan (in cooperation with other relevant institutions); plans the necessary means in advance for the programme evaluation under consideration;<br>• Informs the NIPAC on the Evaluation Plan and the budget available for the contract(s)<br>• Steers the Reference group | | ✔ |

| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| | • Leads the development of the ToR<br>• Manages the contractual arrangements, the budget<br>• Manages the selection and recruitment of the external evaluators<br>• Provides the evaluators with administrative support and required data<br>• Connects the evaluation team with the key evaluation stakeholders and ensures a fully inclusive and transparent approach to the evaluation<br>• Calls for periodic meetings to review progress, to identify solutions<br>• Interacts with evaluators on preliminary and final findings, taking care of avoiding to inadvertently impose interpretations, opinions, and values<br>• Reviews the different deliverables; ensures the final draft report meets quality standards<br>• Ensures the up-take of evaluation findings by national stakeholders<br>• Collects of inputs for relevant meetings and reporting exercises<br>• Organises and coordinates of relevant follow-up | | |
| **Programme Management Units** | Under indirect management, for their actions' portfolio:<br>• Prepare and updates evaluation plans<br>• Contribute to the development of the ToR, by identifying the purpose and scope of the evaluation<br>• Participate in the Inter-service consultative Group/Reference group<br>• Ensures and supports the correct implementation of external evaluations, including:<br>  • Facilitating evaluators' access to stakeholders and officials, and potential and actual beneficiaries of the actions<br>  • Providing updated and reliable data (incl. that of monitoring) to external evaluators promptly and in a usable form.<br>  • Providing feedback on inception report and draft and final evaluation reports.<br>• Ensure the up-take of evaluation findings within the Unit and in the case of generalizable findings, informs relevant colleagues<br>• Provides inputs/comments, as appropriate, for relevant meetings and reporting exercises. | | ✔ |
| **National aid coordinator** | • Approves the evaluation plan, in accordance with IPA II CIR<br>• Oversees implementation of EU support actions<br>• Under decentralised management, coordinates evaluation activities of the different national institutions. This includes ensuring that updated and quality checked data (including that of monitoring) is provided to evaluators promptly and in a usable form.<br>• Allows for decentralised use of monitoring data by organisations that are responsible for collecting and providing data. This can help them understand the relevance and meaning of the data and, therefore, improve data collection, and reduce mistrust in the monitoring system<br>• Ensures the actual up-take and follow-up of evaluation findings at the national level | ✔ | ✔ |
| **Contracting authorities** | • Supervise contractual and financial implementation of EU support actions by means of contacts with the contractors | | ✔ |
| **Line Ministries** | • Supervise implementation of EU support actions, as appropriate<br>• Contribute to internal and external evaluations, as appropriate<br>• Review internal and external evaluation reports, as appropriate<br>• Provide inputs/comments, as appropriate, for relevant meetings and other reporting exercises.<br>• Up-take evaluation findings<br>• Ensure the relevant follow-up | ✔ | ✔ |
| **Other actors** | | | |

| Institution/ Actor | Role | Relevant for ENI | Relevant for IPA II |
|---|---|---|---|
| **Evaluators** | • Respect the principles set up in **BOX 3: EVALUATION PRINCIPLES**<br>• Interact with the Evaluation Manager, with DG NEAR HQ's Monitoring & Evaluation service, with the ISG (in strategic evaluations) or RG, with the stakeholders' group or the expert panel[43] (when relevant)<br>• Draft, present, discuss, and agree with the Evaluation manager and the ISG/RG the evaluation's design that defines how and when they will perform their responsibilities.<br>• Finalise the evaluation questions included in the ToR<br>• Draft the consultation strategy, when relevant<br>• Conduct the evaluation research<br>• Draft, present, and discuss evaluation reports and all other products needed for dissemination—they hold sole responsibility for evaluation findings, which cannot be questioned by any other actor. Evaluators have the responsibility to clearly explain the methods that they follow and their limitations, as well as the limitations of the findings.<br>• They can, if they feel the need to do so, appoint an advisory group, at their own expense using the evaluation's budget, to support them from a methodological point of view. If this is the case, they must present for the ISG/RG endorsement the CVs of the advisory group.<br>• If needed, discuss with the evaluation manager (or report to the ISG/RG or the expert panel) all instances of resistance or of interference. | ✔ | ✔ |
| **Contractors (hiring the evaluators)** | • Provide professional evaluators for whom previous work experiences are carefully checked<br>• Ensure that the evaluators fulfil contract requirements<br>• Manage and supervise the evaluator's work. As such, they are the final responsible actor on the evaluator's work towards the service having launched the evaluation (EC HQ/EUD/national authorities). In this framework, they are expected to:<br>  • Provide training/clear work instructions to the evaluators before the evaluation starts. This should provide guidance to ensure that the evaluation team has a clear understanding of the tasks, of the evaluation process, the content and implications of the different steps.<br>  • Elaborate a task-budget (or similar) for the different evaluation phases, to be monitored according to the final methodological design and to the evaluators' performance. If based on this changes are to be introduced in the evaluation team's composition, they are expected to communicate it as soon as possible to the evaluation manager so that a joint decision is taken.<br>  • Elaborate specific ToR for each evaluator for the different evaluation phases<br>• Allocate sufficient human and financial resources (out of the evaluation budget) to quality control (covering the quality of the process, of the evaluation design, of the evaluators and of the different deliverables) | ✔ | ✔ |
| **Potential and actual beneficiaries** | • Provide updated and reliable data to external evaluators<br>• Provide feedback during the consultation phase, as appropriate<br>• Provide feedback on the draft final evaluation report. | ✔ | ✔ |

*Source: DG NEAR*

[43] Expert panels are not foreseen under the global TOR of the framework contract COM 2015 (to be used in all strategic evaluations (those launched by DG NEAR HQ's Monitoring & Evaluation Service), and can therefore not be organised for evaluations contracted through this channel. The same applies to advisory groups.




# PART 1: LINKING PLANNING/PROGRAMMING, MONITORING AND EVALUATION: M&E THROUGH THE INTERVENTION CYCLE

**BOX 6: SUMMARY OF KEY MESSAGES ADDRESSED IN PART 1**

***The intervention logic articulates the hierarchy of effects that an intervention is expected to produce: from outputs (under direct control), to outcomes (subject to direct influence) and impact (subject to indirect influence). For both outcomes and impact only contribution is to be assessed. As such, the intervention logic of a given intervention needs to be coherent to the broader strategic framework in which the intervention is framed.***

***Its conception is to be considered as a collective endeavour.***

***Monitoring and evaluation are to be thought-through during planning/programming, when the intervention logic of a specific intervention has to be carefully conceived.***

***Indeed, the results' statements included in the intervention logic define what is to be measured, via specific indicators, during the internal and external monitoring and evaluation exercises. Even though additional indicators can be defined in these exercises, the indicators defined during planning/programming must be reported on.***

***The role of the implementing partners in this framework is particularly important.***

***The opportunity framework (i.e. the enabling and hindering factors) in which the intervention occurs is a key element of well-defined intervention logics.***

***It is highly recommended to include intervention logic diagrams in Action documents, as this allows showing the full set of direct and indirect linkages between and within levels, as well as the opportunity framework in which the intervention occurs.***

***Also, it is highly recommended to send the first draft of the intervention logic diagram to both the relevant CoTE(s) and DG NEAR HQ's Monitoring & Evaluation service for a first quality (and informal) check.***

***It is also recommended that an intervention logic diagram is prepared for the relevant country or region (IPA II M/CSP & ENI SSF/CSP), or area/sector. This would ensure the consistency of a given intervention's intervention logic with that of the strategic framework in which it is framed.***

***The same indicators should be used during:***

***Planning/programming as a way to improve the way interventions are drafted (intervention logic's design)***

***Monitoring as a way to steer intervention's implementation (management)***

***According to the Better Regulation, indicators must be RACER: Relevant, Accepted by relevant actors, Credible for non-experts, Easy to report on and Robust, avoiding any manipulation.***

***Indicators must be expressed in neutral terms: they signal a change, not the direction of the change. Finally, their baseline, milestone and target are to be expressed using the same value of measurement.***

***The indicators included in the EU results framework (for ENI) and in the IPA II Performance framework need to be considered, to the extent possible, when defining indicators during planning/programming (and if needed during monitoring and evaluation exercises).***

***Indeed, even if a specific intervention has its own reporting tools and calendar, the EU results framework (for ENI) and in the IPA II Performance framework have been defined at the corporate level to inform on results achieved across relevant partner countries.***

*Source: DG NEAR*

## 8. DEFINING THE INTERVENTION LOGIC AS A STARTING POINT

**While acknowledging that Action documents do not provide for intervention logic diagrams to be included (contrary to log frame matrices), it is highly recommended to include them[44] as these allow showing visually the full set of direct and indirect linkages between and within levels, as well as the opportunity framework in which the intervention occurs.**

### 8.1. What is it?

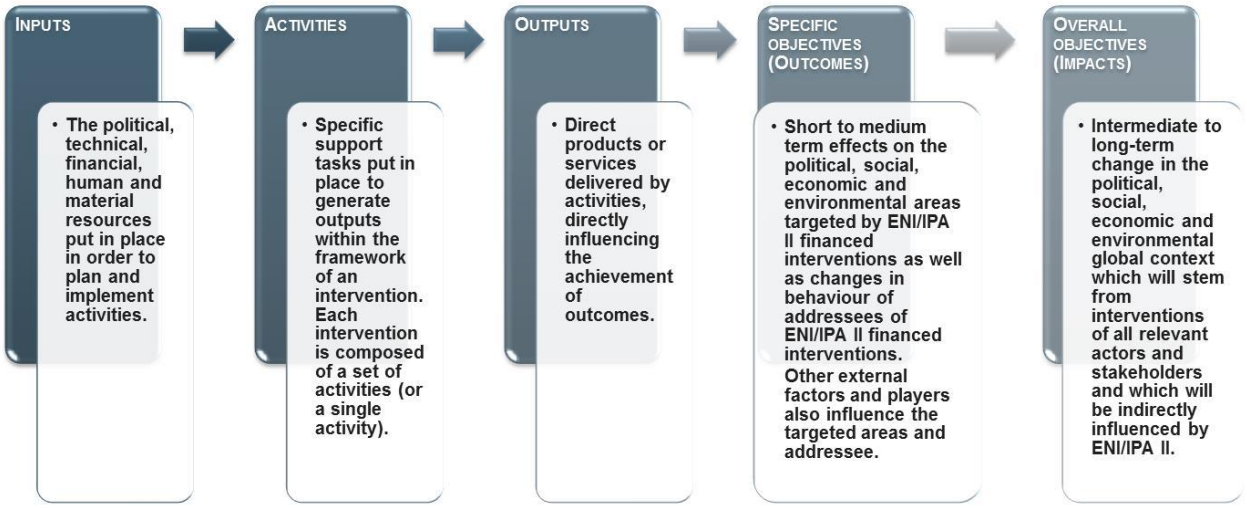
An intervention logic[45] can be defined as the articulated result's chain clarifying the interventions' objectives and translating them into a hierarchy of effects intended to be achieved (up until the level of outputs), directly influenced (outcomes) and indirectly influenced (impacts) by a policy or action.

The rationale behind an intervention logic for a project/programme approach and for budget support is the same. The only difference is that budget support interventions transfer funds to the national treasury of the partner country and aims at improving government policy management and service delivery. This leads to a slightly different result chain, with a different definition of activities (though traditional complementary support is often provided) and the introduction of the concept of induced outputs.

[44] At the end of the Intervention logic chapter for example.

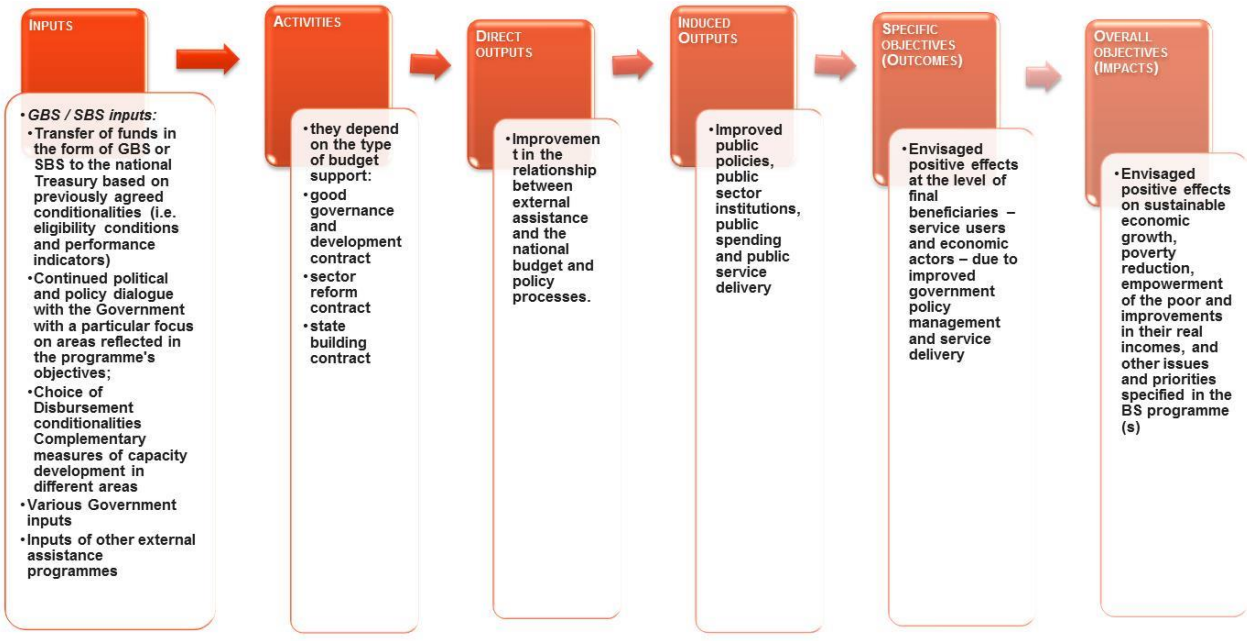
[45] Is there a difference between an intervention logic and a theory of change? A program theory is an explicit theory or model of how an intervention, such as a project, a program, a strategy, an initiative, or a policy, contributes to a chain of results. A programme theory includes a theory of change (which mechanisms produce the desired change) and a theory of action (how the intervention will activate the mechanisms in order to produce its desired change). Funnel and Rogers; other authors, however, use different definitions 2011. Theory of change and intervention logic can therefore be used as interchangeable terms.

**FIGURE 4 : A SIMPLIFIED (LINEAR) DEPICTION OF THE RESULT'S CHAIN OF AN ACTION**




*Source: DG NEAR*

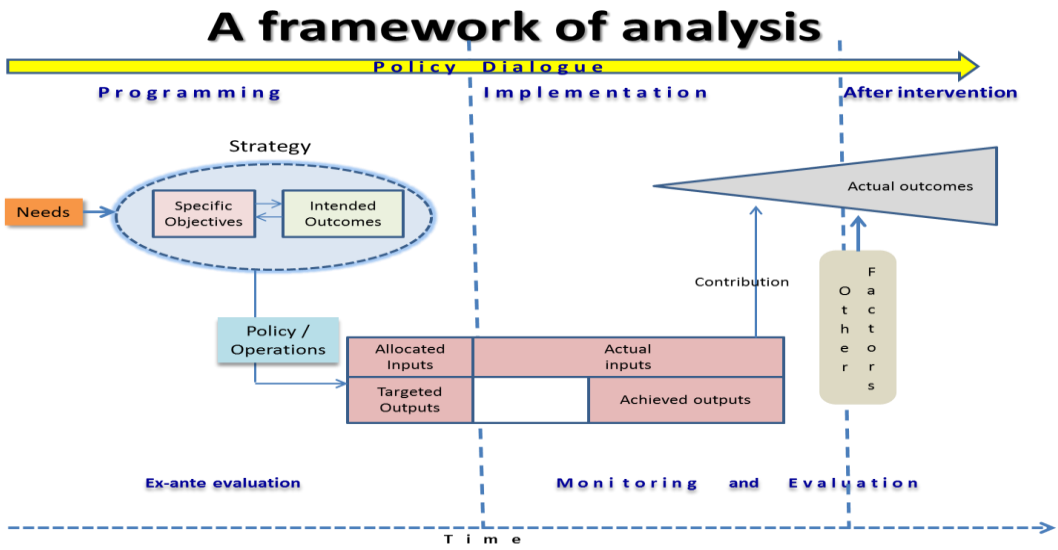
**FIGURE 5 : A SIMPLIFIED (LINEAR) DEPICTION OF THE RESULT'S CHAIN OF A BUDGET SUPPORT PROGRAMMES**




*Source: Budget Support Guidelines Programming, Design and Management A modern approach to Budget support, EC, 2012*

In reality things are more complex. Once the needs are properly assessed a plan can be drawn to define the intended changes pursued and the relevant policies and operations that can serve that purpose. But things normally happen that make the actual strategy pursued, to be discussed with relevant partners, not necessarily identical to the one initially conceived. Moreover, implementation is less linear. Interactions between the different actors (donor agency, national authorities, stakeholders, contractors, in a given context), influence the way in which activities actually take place and outputs are produced. Furthermore, delays take place and other factors influence the way and the timing by which the actual changes materialise, possibly after the end of the intervention (see figure below).

**FIGURE 6 : PLANNING/PROGRAMMING VS M&E: CORRESPONDENCE BETWEEN TERMINOLOGY AND INTERVENTION LOGIC LEVELS**




*Source: The graph represents a re-elaboration of that produced in "Outcome indicators and targets" produced for DG Regional Policy by an expert group led by F. Barca and P. McCann, June 2011*

It is also worth reminding that at the planning/programming stage the intervention logic is usually defined with a backward approach (refer to § **8.3 HOW IS IT DEFINED?**). In evaluation, the intervention logic is defined with an upwards approach, explaining how the interventions' activities are expected to transform inputs into outputs and outputs into outcomes and impacts, through which mechanisms,[46] and if assumptions hold.

**Assumptions** are the necessary and positive conditions that allow for a successful cause-and-effect relationship between different levels of results[47].They are usually less probable at the higher levels of the intervention logic, mainly because they are out of its sphere of control[48]. On the contrary, **risks** are the factors that might hinder the achievement of results. They are also out of the intervention's sphere of control.

**BOX 7: EXAMPLES OF CORE ASSUMPTIONS, RELEVANT FOR ENI & IPA II, TO BE CONSIDERED WHEN DEFINING AN INTERVENTION LOGIC**

Overall, the causal relationships on which the intervention logic is built rely on a set of core assumptions. These assumptions are not related to the features of the context, which may facilitate or hinder the implementation of the IL. The assumptions regard rather the attitude of the partners, their respect of the spirit of partnership and their capacity to translate it into action. Such basic assumptions are highlighted below:

- Common understanding frameworks are built and/or further advanced to address priority issues and mechanisms to further develop them are pursued.
- Obligations under international conventions are implemented and parties cooperate with each other at all levels.
- Interest and commitment of the parties is maintained, adapted and developed throughout the evolution of the context.
- Equality and ownership of partners are respected in ENI/IPA II actions' design and implementation.
- Mobilisation of key stakeholders in partner countries, together with governments, including

[46] "Mechanism refers to the way in which any of the components (of an intervention) or any set of them, or any step or series of steps brings about change. Mechanisms thus explicate the logic of an intervention; they trace the destiny of a programme theory, they pinpoint the ways in which the resources on offer may permeate into the reasoning of the subjects." (Pawson and Tilley, 2004:6).

[47] Handbook on planning, monitoring and evaluating for development results, United Nations Development Programme, 2009.

[48] For a presentation of the spheres of control and direct and indirect influence, please refer to § **8.3 HOW IS IT DEFINED?**

Parliaments, local authorities, civil society and private sector remains a priority.
- The EU ensures enhanced and continued coherence of its multiple external action policies with the ENI/IPA II objectives and interventions.
- Management procedures are transparent, efficient and easy to apply.

*Source: DG NEAR*

As recognised by the Better Regulation, the intervention logic should also consider external factors, which may influence both the performance of the EU intervention or generate the same type of effects[49]. These external factors relate to the context in which the intervention is expected to intervene as well as to the policy and normative framework (refer to **BOX 8: KEY ELEMENTS OF THE OPPORTUNITY FRAMEWORK TO BE CONSIDERED WHEN DEFINING AN INTERVENTION LOGIC**).

**BOX 8: KEY ELEMENTS OF THE OPPORTUNITY FRAMEWORK TO BE CONSIDERED WHEN DEFINING AN INTERVENTION LOGIC**

- **Policy and normative framework** of the relations between the EU, its multilateral partners and partner countries and regions (Treaties, Political Communications, Political Declarations, Joint Communiqués, etc.)
- **EU Internal Policies** (mostly related to Commission Communications, Staff Working Papers, Council Communications and Conclusions) and actions, translating EU interests and priorities at EU and worldwide levels
- Other EU external action instruments (DCI, EIDHR, etc.)
- **EU geographical strategic partnerships** (such as regional and bilateral Agreements with partner countries and regions)
- EU & Partner countries and regions' political and policy dialogues
- EU & Partner countries and regions' political, social, economic and environmental contexts

*Source: DG NEAR*

#### *8.1.1. Is there a difference between an intervention logic and the logical framework approach?*

Their purpose is the same (i.e. to describe how an intervention is expected to lead to results) but the intervention logic (and its diagram) does present some added values compared to the logical framework approach (and its summarised matrix, the log frame; refer to **TABLE 5: TYPICAL STRUCTURE OF A LOG FRAME FOR ENI AND IPA II**). These are:

- As showed in **FIGURE 7: INTERVENTION LOGIC DIAGRAM VS. LOG FRAME MATRIX** a less linear logic and sequence between inputs, activities and objectives (expressed at output, outcome, and impact level). Indeed, contrary to the log frame, the intervention logic helps establishing the direct and indirect linkages (represented with dotted arrows) between effects within the same level and/or between levels. In a log frame matrix, activities are usually directly linked to a single output, which in turn is directly linked to a single outcome.

  In more complex interventions, the link between the initial inputs and the overall objective is non-linear and thus involves more steps in the logical chain, affected by other factors. Complex contexts and interventions require adaptive management,[50] which makes even more important to have synthetic and articulated information on how the program is unfolding.

- Further consideration of the timing issue. Indeed, depending on the intervention at stake, outcomes might happen in the short and/or medium-term, and impacts in an intermediate

[49] European Commission, 2015c: 268.

[50] See O. Barder, "Complexity, Adaptation, and Results", Center for Global Development, September 2012.

and/or long-term. This is clearly reflected in an intervention logic, and usually disregarded in a log frame.

- A better framing of the intervention into the framework and context in which it is expected to intervene. The elements referred to in **BOX 8: KEY ELEMENTS OF THE OPPORTUNITY FRAMEWORK TO BE CONSIDERED WHEN DEFINING AN INTERVENTION LOGIC** are translated into the intervention logic's diagram.
- A better discussion on the theories linking the activities with the intended results.
- The log frame only shows the results' chain, from activities to overall objectives (impact). On the contrary, the intervention logic does include the different set of inputs provided (political, technical, financial and human).

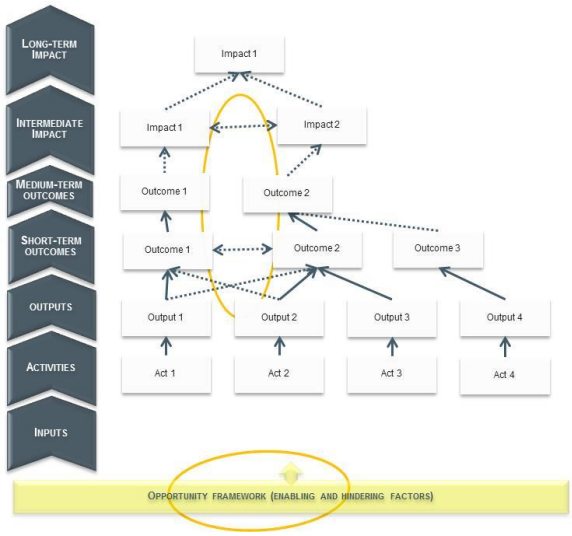
Finally, it is worth mentioning that log frames do present an added value compared to an Intervention logic diagram, which is the inclusion of the indicators for each of the levels. That being said, it is also feasible to do two versions of an intervention logic diagram: one with the results' statements and the other with the corresponding indicators allowing measuring the progress towards the achievement of the results.

**TABLE 5: TYPICAL STRUCTURE OF A LOG FRAME FOR ENI AND IPA II**

| Intervention | Indicators | Source of verification | Assumptions |
|---|---|---|---|
| **Overall objective** (Impact) | How will the overall objectives be measured, including quantity, quality, and time? | How will the information be collected, when and by whom? | |
| **Specific objective** (Outcome) | How will the specific objectives be measured, including quantity, quality, and time? | As above | If the specific objective is achieved, what assumptions must hold true to achieve the overall objective? |
| **Results** (Outputs) | How the results are to be measured including Quantity, Quality, and Time? | As above | If Results are achieved, what assumptions must hold true to deliver the specific objective? |
| **Activities** | | | If Activities are completed, what assumptions must hold true to deliver the results? |

*Source: DG NEAR*

**FIGURE 7: INTERVENTION LOGIC DIAGRAM VS. LOG FRAME MATRIX**




| Log frame for Project/programme | |
|---|---|
| Impact | Impact 1<br>Impact 2 |
| Outcomes | Outcome 1<br>Outcome 2<br>Outcome 3 |
| Outputs | Output 1.1<br>Output 2.1<br>Output 2.2<br>Output 3.1 |
| Activities | Activity 1.1.1<br>Activity 2.1.1<br>Activity 2.2.1<br>Activity 3.1.1 |

| Log frame for Budget support programme | |
|---|---|
| Impact | Impact 1<br>Impact 2 |
| Outcomes | Outcome 1<br>Outcome 2<br>Outcome 3 |
| Induced Outputs | Output 1.1<br>Output 2.1<br>Output 2.2<br>Output 3.1 |
| Direct Outputs | Output 1.1.1<br>Output 2.1.1<br>Output 2.2.1<br>Output 3.1.1 |

*Source: DG NEAR*

**As mentioned earlier, intervention logic diagrams allow showing visually the full set of direct and indirect linkages between and within levels, as well as the opportunity framework in which the**

**intervention occurs. Without an intervention logic diagram, the orange ovals of FIGURE 7: INTERVENTION LOGIC DIAGRAM VS. LOG FRAME MATRIX would remain overlooked.**

### 8.2. Why do we need it?

The intervention logic is a crucial element in planning/programming, monitoring and evaluation.

- **It constitutes the basis for the identification of suitable indicators**[51] against which the intervention will be monitored and evaluated. Each result's statement included in each of the levels of the intervention logic is to be accompanied by one or more indicators.
- As mentioned above, the indicators defined up to the level of the outputs, that determine the sphere of control of the interventions' achievements, will be used in monitoring. In certain cases, mainly depending on when the monitoring exercise is launched, contribution to the outcomes will also be part of monitoring and for this to happen indicators defined during planning/programming at the outcome level will be the ones for which a value of measurement will need to be provided.
- In evaluation, the intervention logic will be the basis for the definition of the evaluation questions[52]. If the intervention logic is not made explicit and/or not sufficiently articulated in the planning/programming document, it will need to be made explicit and reconstructed during an evaluation exercise, thus requiring a fine tuning of the evaluation questions during the inception stage.
- As mentioned above, evaluations do mainly focus on the spheres of direct (outcomes) and indirect (impacts) influence. As such, indicators defined for these levels of the intervention logic will be used in evaluation. It must be highlighted though that depending on the specific purpose and scope of the evaluation exercise additional indicators can be defined.

### 8.3. How is it defined?

#### *8.3.1. Pre-conditions*

First of all, it is recommended to use a participatory approach by means of a group discussion or similar among key stakeholders (DG NEAR/EUD programme manager, representative of the national/regional authorities and of direct addressees, if these do not correspond to the authorities). In the case of planning/programming (joint programming for ENI), or in the case of budget support[53], this participatory approach is requested by definition; indeed, the definition of a common result framework (for ENI) and of a Performance Framework (for IPA II) is of utmost importance here.

In the framework of DG NEAR actions, a participatory design has proven to be a key success factor for actions. Indeed, the most recent ROM results for ENI[54]&[55] showed that in cases in which this collaborative

[51] Refer to § **9 DEFINING INDICATORS: SETTING THE BASIS FOR M&E** for further details.

[52] Refer to **BOX 51: ELEMENTS OF AN EVALUATION QUESTION** for further details.

[53] This document provides common definitions and guidelines to develop an intervention logic and to identify and design related indicators; It is thus to be considered as a complementary tool to DG EuropeAid budget support guidelines (https://encrypted.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&cad=rja&uact=8&ved=0ahUKEwjv6ce48YHNAhXEVxQKHaA6BOwQFgg1MAQ&url=http%3A%2F%2Fwww.respaweb.eu%2Fdownload%2Fdoc%2FProgramme%2520of%2520Workshop%2520on%2520Sector%2520Budget%2520Support.pdf%2F75d491c286c54bee19068dfa7eae7834.pdf&usg=AFQjCNFt_5OYPDwD2J6SdMkEFCUSHCK-WA) and to former DG enlargement sector budget support guidelines (https://encrypted.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&uact=8&ved=0ahUKEwjv6ce48YHNAhXEVxQKHaA6BOwQFggoMAI&url=http%3A%2F%2Fwww.evropa.gov.rs%2FDocuments%2FHome%2FDACU%2F5%2F28%2F247%2FELARG%2520SBS%2520Executive%2520Summary.doc.pdf&usg=AFQjCNF5-BozwGVXUQhYLczacU9gss6VZg).

[54] A total of 144 actions were monitored in 2015 for ENI South and East.

[55] ROM synthesis results for IPA II countries were not available at the time of drafting of the present document.

approach was followed, it resulted in: i) an improved tailoring of services and activities, ii) an enhanced ownership and commitment from both EU and partner countries, iii) an increased active leadership by the actions' local counterparts and iv) positive repercussions on performance under all DAC criteria. More generally speaking, quality in programming has also proven to have a positive impact on efficiency and effectiveness.

In this framework, the sequential and temporal path of an intervention logic needs to be conferred an important place during planning/programming discussions. Indeed, each element of a results' chain is expected to be achieved in different moments in time. This means that the more time passes, the more indirect is the sphere of influence of a single intervention. As such, the expected contribution of an intervention in these different settings and moments, each one affected by different addressees, is to be kept in mind when devising the intervention's rationale.

**BOX 9: AN INTERVENTION LOGIC'S SPHERES OF CONTROL, DIRECT AND INDIRECT INFLUENCE**

**Sphere of control:**

Inputs, activities and outputs within the intervention's own sphere of control.

**Sphere of Direct influence:**

Short and medium term outcomes in terms of the engagement, take-up (use), and actions of organizations, institutions, communities and individuals who are directly 'in touch' with the intervention.

In the case of Budget support, the sphere of direct influence only covers the induced outputs.

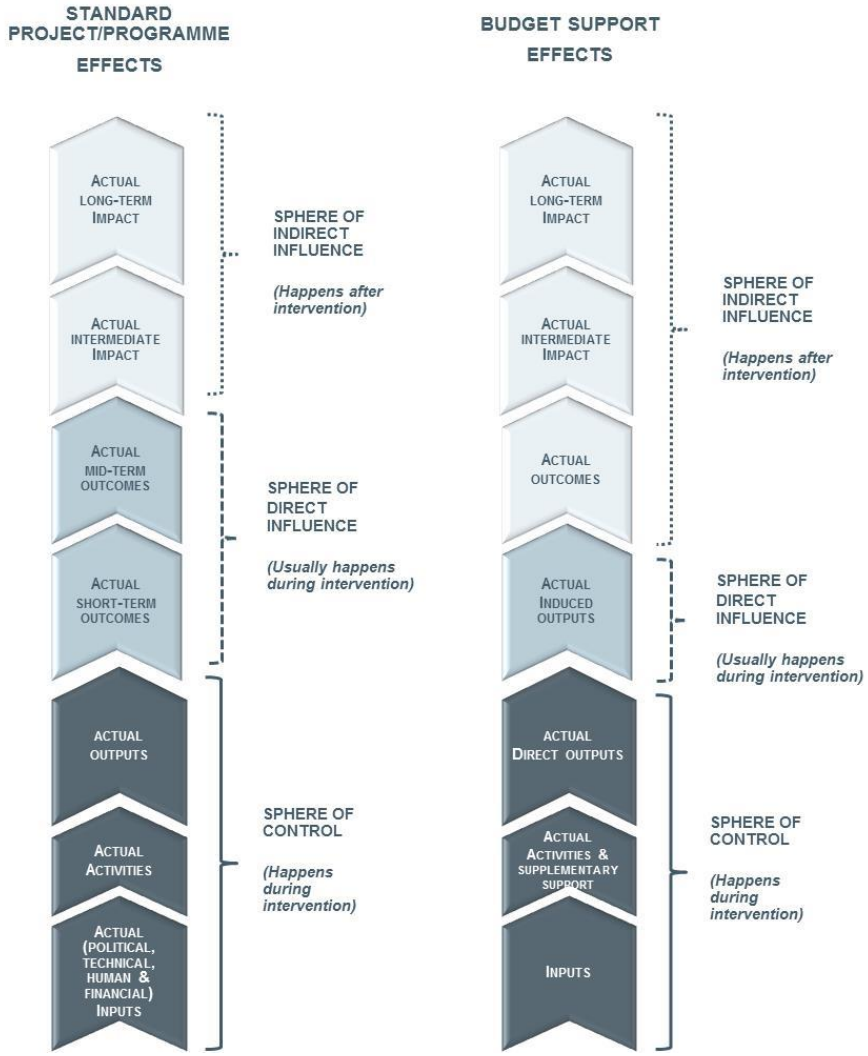
**Sphere of Indirect influence:**

The intermediate and long term desired impacts that happen in the political, social, economic and environmental global context and that require the involvement of addressees that were not directly 'in touch' with the intervention.

In the case of Budget support, the sphere of indirect influence covers the outcomes and the intermediate and long term desired impacts.

*Source: DG NEAR, adapted from The Need to Build Reach into Results Logic and Performance Frameworks, Steve Montague, Nancy Porteous and Sanjeev Sridharan, January 2011*

**FIGURE 8 : SPHERES OF CONTROL, DIRECT AND INDIRECT INFLUENCE: PROJECT/PROGRAMME APPROACH VS. BUDGET SUPPORT**




*Source: DG NEAR*

Finally, **it is highly advisable to involve in this process a member of DG NEAR HQ's Monitoring & Evaluation service**, even via email, at least once a first draft is available. Indeed, country/region/sectorial knowledge is key and the programme manager and national/regional stakeholders would offer these, but the evaluative thinking[56] is equally important. This support is not to be understood as a formalised step of the existing quality review process; this support is offered to colleagues upon request on an informal basis and it would have the benefit of facilitating and probably even shorten the formal quality review process.

[56] Evaluative thinking is an increasingly important topic in the field of evaluation, but a common or recognized definition does not exist yet. For the sake of clarity, we can use the following definition given by Michael Quinn Patton: Evaluative thinking involves systematic results-oriented thinking about what results are expected, how results can be achieved, what evidence is needed to inform future actions and judgments, and how results can be improved in the future. Refer to http://www.theclearinitiative.org/EvaluativeThinkingReport_FINAL_online.pdf.

### *8.3.2. Steps: defining the intervention logic's diagram (steps 1 to 5), developing the intervention logic's narrative (step 6)* [57 & 58 & 59]

Before drafting any narrative, the intervention logic of the intervention needs to be translated into an effects diagram. This is a simple and crucial way to show the soundness of the result's chain. For this to happen, the following steps are needed:

#### 8.3.2.1. Step 1: identification of needs (the overall objective(s))

**Step 1**: The starting point in setting up intervention logics is **to identify the needs to be satisfied**[60]. These needs are translated into the overall objective(s) of your intervention (your intended impacts).

These objectives are intermediate and long-term changes in the political, social, economic and environmental context that are expected to be indirectly influenced by your intervention. Indeed, these changes will stem from interventions of all relevant actors and stakeholders, not only by your intervention.

Examples of overall objectives (Impacts):

Long-term:

- Attainment of internationally agreed goals, incl. sustainable growth and sustainable management of rapid urbanization
- Sustainable economic, social and environmental development
- Achievement of inclusive democracy

Intermediate term:

- Increased resilience to climate change (global warming limited to 2*C above pre-industrial levels)
- Increased citizens' scrutiny and participation in public life, incl. in enhanced and sustainable local economic development
- Local and regional governments are accountable and responsive to citizen demands for access to basic services
- Social cohesion strengthened
- Greener growth.

As showed in **BOX 12: REPRESENTING A SIMPLE PROGRAMME THEORY FOR TRAINING** and in **FIGURE 9: LINKING THE INTERVENTION LOGIC OF AN INTERVENTION WITH (ENI & IPA II) STRATEGIC FRAMEWORK**, it is important to highlight that the overall objectives of a specific intervention (and consequently the lower levels as well) need to be coherent and feed into ENI and IPA II strategic frameworks. In other words: the intervention logic of a single intervention needs to be nested into the intervention logic of a country/region/sectorial strategy. Indeed, the contribution of a single intervention to the results that are expected from a country/region/sectorial strategy is only one among others, and its actual reach is also therefore more limited. The overall interventions' objectives are

[57] Refer to Annex 2 for a template of an action's intervention logic (both in English and French).

[58] Refer to Annex 3 for a template of a budget support programme intervention logic (both in English and French).

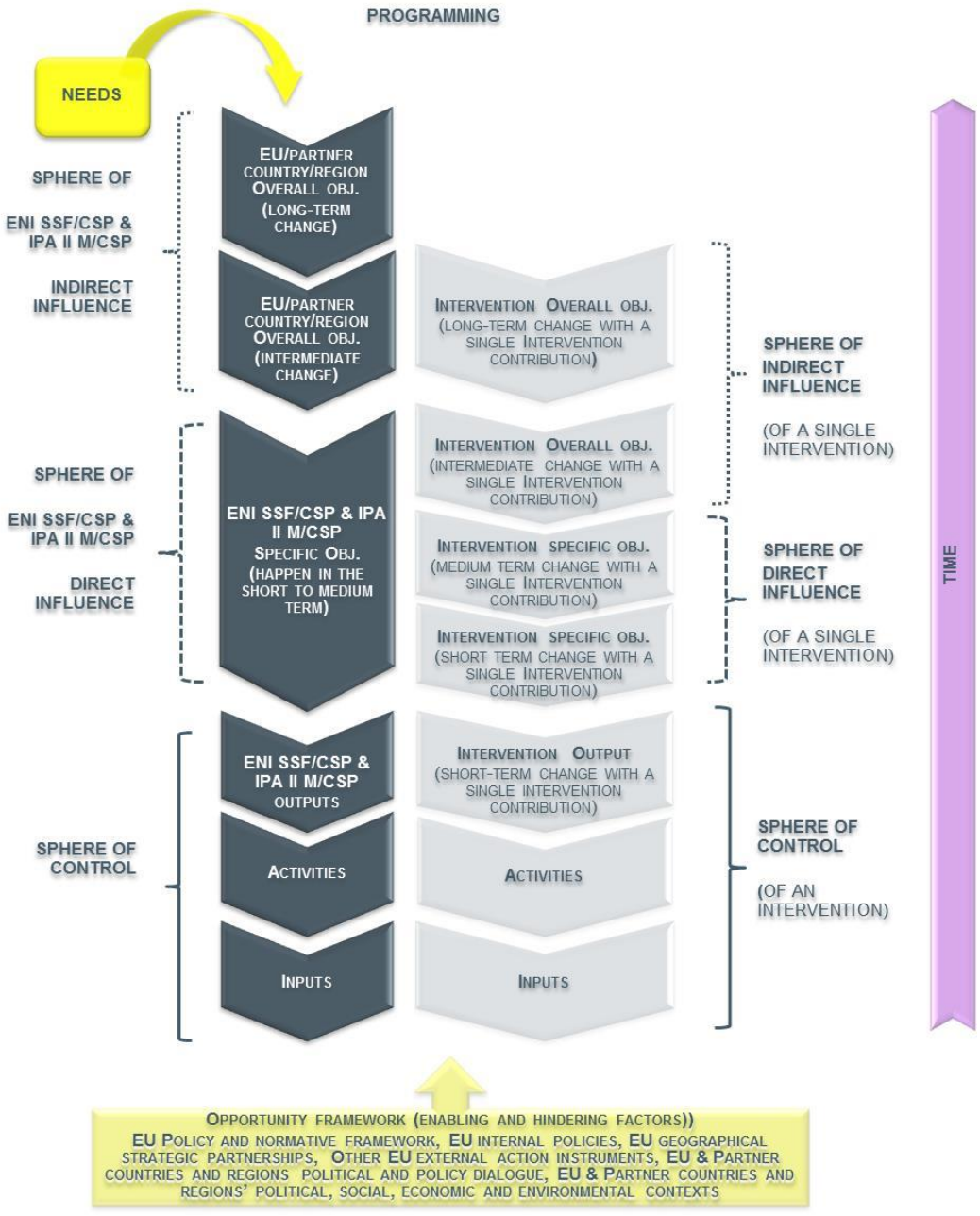
[59] An example of a DG NEAR relevant intervention logics is provided in Vol 1, p. 3 of the Strategic evaluation of the EU Cooperation with the Hashemite Kingdom of Jordan (2007-2013): http://ec.europa.eu/europeaid/strategic-evaluation-eu-cooperation-hashemite-kingdom-jordan-2007-2013_en. For a French version, another example can be found in the Évaluation de la Coopération de la Commission Européenne avec la Tunisie, in Vol 2, p. 59 at: http://ec.europa.eu/europeaid/strategic-evaluation-eu-cooperation-tunisia-1995-2008_en.

[60] A **need** is defined as "something without which unsatisfactory functioning occurs".

usually expected to be achieved in a timespan that goes from mid to long term. Using a nested log frame approach, it can be decided to situate the overall long-term objective(s) (impact) of an intervention at the level of the intermediate action programmes objectives.

**In order to facilitate and ensure the consistency of a given intervention's intervention logic with that of the strategic framework in which it is framed, it is highly recommended that at HQ's Unit/EUD level, the intervention logic diagram of the relevant country or region (IPA II M/CSP and ENI SSF/CSP), or area/sector is done.** Here again, it is advisable to involve in this process a member of DG NEAR HQ's Monitoring & Evaluation service

**FIGURE 9: LINKING THE INTERVENTION LOGIC OF AN INTERVENTION WITH (ENI & IPA II) STRATEGIC FRAMEWORKS**




*Source: DG NEAR*

**BOX 10: FROM ENI/IPA II-RELATED OVERALL POLITICAL OBJECTIVES AND PRIORITIES TO ACTIONS OVERALL OBJECTIVES**

The overall EU and partner countries objectives and priorities are defined for the Neighbourhood in the European Neighbourhood Policy and for enlargement countries in the Annual Enlargement package. These are then translated into Single Support Framework, bilateral ENP Action Plans or Country Strategy Papers and related multi-annual indicative programme[61] for the Neighbourhood and into country/multi-country (regional) strategy papers and related annual/multi-annual action programmes for enlargement countries.

Specific actions need to be conceived in a way that they contribute to the achievement of the objectives defined under their relevant strategic framework.

In the context of IPA II, there are linkages between the intervention logic of the Indicative Country/Multi-Country Strategy Paper (M/CSP) and that of the Sector Planning Document. M/CSP specific objectives should ideally be at the same level as the overall objectives of the Sector Programme/Plan, whilst CSP results would be at the level of the specific objectives of the Sector Programme/Plan. The same logic would apply to the linkage between the Sector Programme/Plan and the individual actions, according to the principle of the interlocking or nested Logical Framework.

*Source: DG NEAR*

#### 8.3.2.2. Step 2: identification of the specific objectives (the outcomes)

**Step 2**: With a backward approach, **identify the outcomes (your specific objectives) that are needed for the overall objective(s) to be achieved**. The assumptions leading from outcomes to impacts need to be discussed.

These outcomes are short to medium term changes on the political, social, economic and environmental areas targeted by your intervention, as well as changes in behaviour of addressees of your intervention. These changes are expected to be directly influenced by your intervention, but it is important to recall that other external factors and players also influence the targeted areas and addressees of your intervention. When defining this level, it is particularly important to clearly identify and precise the target group(s).

**They need to be formulated as a result statement**.

The first question to be asked relates to **describing which improvements are expected for whom**.

**BOX 11: DESCRIBING OUTCOMES**

We use the following example on training for public servants, aimed at introducing improvements in public sector performance

What is the definition of "improvements" for that particular area of public work? Programme managers can use this opportunity to go into what "higher quality of existing services" means (services by whom, for whom, when) and, of course, how it is measured. The description of the improvements should be as concrete and precise as possible, and include all relevant dimensions

For example: the new procedures or services are in place, achieved higher quality of existing services for firms or citizens, more gender equality in the access to services, better fit between public sector's supply and citizens' demand for services, reduction of the time it takes for citizens to obtain services from the public sector, etc.

*Source: DG NEAR*

Exploring the improvements expected also means to face **which groups benefit and which groups do not**, and how equilibriums change. This is useful for implementation, because it helps **anticipate where resistances may arise**. For example, requiring great changes in work organisation of public workers organised in a powerful trade union) may prove difficult: they have the power to both oppose the change

[61] For the few countries with no ENP Action Plan, Association Agenda or equivalent documents (currently Belarus, Syria and Libya, Algeria).

openly and to resist it in their everyday work. On the contrary, benefits accruing to a social category which is not organised (e.g., dwellers in remote internal areas or young women job seekers) or to a group which is socially weak (e.g., minority children or transnational refugees) will not elicit strong support. Identifying these sources might help devise strategies to gather support and to decrease resistances.

Often, **there are multiple dimensions to the improvements expected**, rather than just one: among these, only one or two will be measured. While the first step relates to describing the expected outcome, the second step relates to identifying the most relevant dimension.

Examples of specific objectives (outcomes):

- Increased EU influence in partner countries on policy formulation with regard to bilateral or global issues of global concern, in line with EU interests
- Opportunities for low carbon and environmental friendly EU investment, technology and green goods/services (green trade) improved
- Improved business environment in partner countries notably for EU and partner countries SMEs
- More sustainable energy policies in strategic partner countries, in line with EU interests
- Demand-driven decentralized cooperation schemes adopted and implemented
- Appropriate territorial development strategies and policies for better service delivery at the local level adopted

**BOX 12: REPRESENTING A SIMPLE PROGRAMME THEORY FOR TRAINING** [62]

Once the concept of improvement in public sector performance has been fully explored, it is necessary to reconstruct the causal chain leading from training to that particular bunch of improvements at the outcome level. Training, in fact, might not be the best (most expedient, fastest, or most economical) way of bringing about the outcomes—it might even not be conducive to the expected improvements at all, for example whenever the obstacle is lack of physical equipment (if the equipment is not there), training employees in IT use might just engender frustration or whenever there are conflicting incentives (i.e. people know how to do things, but do not because it is not in their interest). Therefore, the work starts with asking why the activity (training), leading to an increased knowledge in IT (output) should work.

A traditional linear theory identifies the assumptions, which have to hold for the expected outcomes to materialise (e.g., trainees must remain in the organisation). A realist theory identifies different groups of trainees (group A, integrated organisation members; group B, the disenfranchised; and group C, the up-and-rising professionals) and associates different outcomes to each context.

Training is delivered to public servants: different contexts will activate different mechanisms engendering various types of learning. At each step, it is necessary to ask which change we expect and which unexpected change might materialise, how to capture this change, and which change may be detected routinely and which can be detected through other activities (such as evaluations, ad hoc surveys, or on site visits).

*Source: DG NEAR*

#### 8.3.2.3. Step 3: identification of the outputs

**Step 3**: Again, with a backward approach, **identify the outputs** (called Results in ENI/IPA II planning/programming templates) **that are needed for the outcomes to be achieved**.

**They need to be formulated as a result statement.**

[62] Refer to Annex 4 for an example of a theory of change for a training action.

**BOX 13: EXAMPLES OF TYPOLOGIES OF OUTPUTS**

**Typology 1: Strengthened individual/institutional knowledge and capacities**

- New/strengthened sectorial/ thematic knowledge and capacities

**Typology 2: Improved intelligence**

- Improved policy evidence for public and private stakeholders
- Improved evidence for the development of common agendas

**Typology 3: Strengthened, better informed and more effective relations**

- More effective peer-to-peer ties
- Debate relevant to EU and joint agendas is activated/reinvigorated
- Improved bases (i.e.; ownership, convergence of ideas, etc.) for (new/on-going) policy dialogue and better coordination

**Typology 4: Better advocacy and improved awareness**

- Increased advocacy concerning EU and EU and partner countries interests
- Improved networking, advocacy and outreach
- Strengthened dialogue/engagement with partner country public/target groups on i) EU policy priorities and actions, business interests, etc. and/or ii) EU and partner countries' common areas of interest

*Source: DG NEAR*

As mentioned under § **8.1 WHAT IS IT?**, budget support programmes distinguish between direct and induced outputs.

**BOX 14: EXAMPLES OF DIRECT OUTPUTS IN BUDGET SUPPORT PROGRAMMES**

**Direct outputs of GBS/SBS:**

- Increased size and share of external assistance funds made available through the national budget.
- Increased size and share of budget available for discretionary spending.
- Increased predictability of the disbursement of external funds.
- Policy dialogue, conditionalities and TA/capacity building activities better coordinated and more conducive for implementation of government strategies.
- External assistance as a whole (including budget support) better harmonised and aligned to government policies and systems.
- Reduced transaction costs of providing aid

**Other effects by various Government inputs:**

- Domestic revenue funding and domestic policy inputs

**Other effects by other external assistance**

*Source: Budget Support Guidelines Programming, Design and Management, A modern approach to Budget support, EC, 2012*

**BOX 15: EXAMPLES OF INDUCED OUTPUTS IN BUDGET SUPPORT PROGRAMMES**

- Improved macroeconomic and budget management (such as fiscal, monetary, trade and economic growth policies).
- Increased quantity and quality of goods and services provided by the public sector
- Strengthened PFM and procurement systems (transparency, fiscal discipline, oversight, allocative and operational efficiency)
- Improved public policy formulation and execution processes
- Strengthened public sector institutions.
- Strengthened links between the Government and oversight bodies in terms of policy formulation and

approval, financial and non-financial accountability and budget scrutiny

- Other improvements in governance issues (e.g. enhanced decentralisation, application of rule of law, human rights

*Source: Budget Support Guidelines Programming, Design and Management a modern approach to Budget support, EC, 2012*

#### 8.3.2.4. Step 4: identification of the activities

**Step 4**: Identify the activities that are needed for the outputs to be produced.

**BOX 16: EXAMPLES OF TYPOLOGIES OF ACTIVITIES**

**Typology 1: Training, knowledge creation and other formal events**

(Formal training of, and knowledge creation among, individuals, plus the organisation of formal group events such as conferences, seminars, debates, workshops)

**Typology 2: Knowledge products**

(Preparation of technical products (studies, reports, strategic plans, roadmaps, tools, databases, provision of information, examples of best practice, monitoring reports, publications))

**Typology 3: Expertise, exchanges and partnerships**

(Organisation of visits, exchanges, business missions, study tours, provision of expertise)

**Typology 4: Promotion, outreach and engagement**

(Promoting EU and EU and partner countries joint policy priorities, (campaigns, communication events, press releases, etc.) actions, business interests, and raising the visibility and awareness of the EU and partner countries)

*Source: DG NEAR*

**BOX 17: EXAMPLES OF ACTIVITIES IN BUDGET SUPPORT PROGRAMMES**

- transfer of €X M over the period (specify partner country fiscal years over which disbursements expected);
- continued political and policy dialogue with the Government with a particular focus on areas reflected in the programmes objectives, as well as wider issues concerning the country's commitment to the fundamental values (in the case of Good governance and development and State building contracts);
- continued dialogue between the EU Delegation and other donors to coordinate and further align our development cooperation with a view to avoiding duplication of activities and relieving the Government from multiple reporting duties;
- regular monitoring of budget support eligibility criteria
- Complementary support (Activities leading to reinforcing Government's capacities, works, supplies, etc.)

*Source: Budget Support Guidelines Programming, Design and Management, a modern approach to Budget support, EC, 2012*

#### 8.3.2.5. Step 5: develop the indicators

**Step 5:** Develop the indicators, with associated baselines, milestones and targets. The process of identifying indicators can be considered as a test of the soundness of the interventions' objectives and can lead to improvements in design. [63] Also, if the previous steps are soundly done, this step should easily flow.

**While acknowledging that the present guidelines do not aim at defining the quality review process, it is highly recommended to send the intervention logic's diagram to quality review (to both relevant CoTEs and DG NEAR HQ's Monitoring & Evaluation service) once steps 1 to 5 are finalised.**

[63] Refer to § **9 DEFINING INDICATORS: SETTING THE BASIS FOR M&E** for further details.

#### 8.3.2.6. Step 6: develop the narrative

**Step 6**: Develop the narrative. The intervention logic must be explicit and soundly articulated in the Action documents. The effects hierarchy must be framed, as mentioned earlier, in the context and policy framework in which the intervention will develop. Assumptions and risks (and related mitigating measures) need to be discussed here as well.

**BOX 18: EXAMPLES OF SUPPORT QUESTIONS AND TIPS IN ELABORATING/RECONSTRUCTING THE INTERVENTION LOGIC**

The following set of **questions** provides orientation in identifying the:

Expected change

- *How do you describe the expected change? (For example, how much time elapses between the request for a service and its delivery?*
- *Who will benefit from the change? How do women benefit from the activities? How will existing equilibriums change? (E.g., how likely is it for men to feel their position has worsened in relative terms to women's? If the expected outcome is an improvement in travel time on a railway between two points, how would this affect bus companies?).*
- *Which gender differences are there in the access to the services? How does each main stakeholder describe the expected change?*
- *Which differences in describing the improvement emerged in intervention development or in negotiations?*

Contexts

- *Who are the focused/targeted groups in society?*
- *Which subgroups exist among the potential beneficiaries? How will they react?*

Mechanisms and outcomes

- *Which organisations will have to conduct which activities? Which human and organisational resources do implementers need in order to activate the financial resources within the set time?*
- *What evidence is there that the outputs actually are translated into the outcomes expected? How long does it take for intervention beneficiaries to actually start benefiting from the outputs? Which is the sequence in which this happens?*
- *Which mechanisms presumably operate in order to translate outputs into outcomes?[64] How do mechanisms operate differently for women and for men?*

**TIPS**

**Always make sure the answers to the questions are in the active form**, since this makes clear who acts and who reaps benefits. This is important both when describing the expected change and the groups affected and when describing the organisations which have to act:

- who receives the benefits
- who has the responsibility to act

Where to look for the **information**?

- National reporting systems
- Ex post evaluations of similar actions in the past in the same place
- Research in the sector of interest
- Sector expertise
- Ex ante evaluations

*Source: DG NEAR*

[64] Evaluations are concerned with the task of verifying which mechanisms have operated in reality. For the construction of an indicator system, however, it suffices to hypothesise which mechanisms can be at play, on the basis of existing evidence.

# 9. DEFINING INDICATORS: SETTING THE BASIS FOR M&E

## 9.1. Definition, purpose and uses

### *9.1.1. What is an indicator?*

An indicator is a 'quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of an actor'[65].

A performance indicator is "a variable that allows the verification of changes in the intervention or shows results relative to what was planned"[66]. Performance indicators help build systematic knowledge on actions, policies etc. degree of success in carrying out the planned activities, producing planned outputs and contributing to intended outcome and impacts on the political, social, economic and environmental contexts. Information from an indicator system answers these questions:

- *Which of the desired changes has materialised?*
- *How far are we from the target?*

The indicators system is **focused on *desired* change -** the positive/negative change that programmers and stakeholders expect. The indicators cannot, per se, demonstrate that the change is actually attributable to the intervention, but they **show how far we have improved and, ideally, how far we are from the target**.

### *9.1.2. Why do we need indicators? Their purpose and use during planning/programming, monitoring and evaluation*

The main purpose is "to give interested stakeholders the possibility of verifying if, and to which extent, by means of policy/*action* implementation, the expected outcomes and impact are going to be (or have been) achieved and provide therefore evidence to support a possible change".[67]

**Indicators cater to the needs of a specific category of stakeholders**: European Commission top decision-makers, European Commission Headquarters, Delegations, beneficiary countries top decision-makers, beneficiary countries coordinating authorities and implementing agencies, groups affected by activities, and the general public. Each of the intervention actors has their own responsibilities, their own areas of decision-making and their own information needs. As a result, not all indicators are useful at all levels. On the contrary, it is generally accepted that each actor requires an operating report with a small number of indicators, selected as the most relevant in relation to the nature of the decisions that have to be made.

As such, putting in place effective and consistent indicators systems provides a powerful tool capable of **focusing planners', programmers', and implementers' attention on the desired change**. Indicator systems focus on expected effects. They can alter implementers' behaviour in both positive and negative ways. They help identify areas where knowledge is missing and should be collected through evaluations or other studies. However, at the outcome and impact levels, they cannot demonstrate that change is the sole consequence of intervention activities. It is necessary to launch evaluations in order to understand which unexpected effects have materialised, to grasp whether there have been undesirable distortions in behaviour, and to ascertain the relationship between the intervention and its effects. Identifying sources for the necessary data prompts improvements in statistical systems, and requires a deep rethinking of the way administrative data is accumulated, made available, and preserved.

[65] OECD DAC, 2010a: 25.

[66] Ibid., 29.

[67] DG NEAR's Note on indicators.

Indicator systems are therefore used in programming, monitoring and evaluation (and related reporting mechanisms).

9.1.2.1. During planning/programming: a way to improve the way interventions are drafted (intervention logic's design)

As mentioned under 8.3.2, the definition of indicators is a crucial part of planning/programming. Once the different levels of the intervention logic are defined (from overall objectives to activities), indicators for each of them are to be identified. They therefore allow further clarifying, detail and focusing the intervention logic by defining concrete targets in terms of values, timing and addressees, resulting in **improvements in the way interventions are conceived and drafted**. Indeed, identifying indicators requires a close scrutiny of interventions (a purpose linked to **learning** as well).

9.1.2.2. During monitoring: a way to steer intervention's implementation (management)

Information from a performance indicator system may be used for **management**: to **steer policies, instruments, actions,** etc. by providing an idea of whether implementation is on course in order for intervention staff, programme managers, and policy makers to devise and implement correction measures or to continue as planned.

This is crucial in a period of tightening budgets and search for efficiency, and implies that there is a founded theory of the relationship between the change in the indicator, the changes in the phenomenon of interest, and the changes which policy makers, programme managers, and intervention staff can effect in interventions—in short, that there is a well-founded theory of change.

NEAR MIS is worth mentioning here as it now aims at reporting on a selection of key performance indicators.

9.1.2.3. During evaluation: a way to demonstrate results to internal and external stakeholders (accountability)

The knowledge accumulated through an indicator system may **demonstrate whether the desired change is happening:** it can be one of the tools to use in order to respond to requests of **accountability**. Potentially, the **audience for this knowledge is both internal** (higher management or control and audit authorities) **and external** (e.g., policy makers, implementing agencies, and public opinion both within Member States and partner countries).

In this framework, access to indicator data should be given to all the partner country organisations that fill in data in the indicator system. In order to complete the availability of information for the general public in both Member States and partner countries, data from the performance system could be disclosed as **open data** in order to inform the public and allow for creative utilisation of the data.

This purpose calls for indicator systems to provide aggregated data on measures across interventions and countries. Albeit important, this is possible solely for the portions of interventions which share the same goals and activate similar mechanisms and only insofar as measures are homogeneous. This implies that indicators are defined in the same way and that the raw data is collected precisely in the same way.

## 9.2. Type, components and characteristics of indicators

### *9.2.1. Type of indicators*

9.2.1.1. Quantitative versus qualitative indicators

A balanced indicator system includes both quantitative and qualitative indicators.

Quantitative indicators refer to dimensions of a phenomenon, which can be easily counted. They are expressed in numerical form—absolute numbers, percentages, rate or ratio.

**BOX 19: EXAMPLES OF QUANTITATIVE INDICATORS**

**Overall objective (Impact) level**:

- Number of pollutants in water bodies
- Number of recorded illegal border crossings
- GDP per capita growth (annual %)
- Share of population by educational attainment
- Renewable energy share in the total final energy consumption
- Share of older persons receiving pensions
- Ratio of the youth population (aged 15-24) to the working-age population (15-64)

**Specific objective (Outcome) level**:

- Number of people with access to justice services
- Number of SMEs applying Sustainable Consumption and Production practices with EU support
- Number of Legislative reviews related to co-operation in border management conducted
- Ratio of new laws amended within one year of their adoption.
- Number of International Labour Organization Conventions ratified, by type of convention
- Share of social assistance beneficiaries registered as unemployed involved in active labour market programs
- SMEs participation share in energy efficiency
- Number of national and local disaster risk reduction strategies

**Output level**:

- Number of outputs produced (e.g. structures, plans, systems, proposals, reports etc.) for improved service provision
- Number of research and knowledge products drafted referenced in an institution's advocacy strategy
- Number of drafted proposals to improve an institution's governance mechanisms
- Percentage of women among public employees receiving training
- Number of people receiving rural advisory services with EU support
- Number of participants who have enhanced their skills/ knowledge/ awareness of the sector under consideration.
- Number of recommendations, agreements, conclusions, etc. emanating from conferences/ seminars/meetings/workshops/ policy dialogues/events
- Number whose opinions change

*Source: DG NEAR*

Qualitative indicators measure quality, opinions, perceptions, stages in a process, or a status. They can be expressed in various forms: yes/no, compliance with, quality of, extent of, level of. They can also be expressed in numerical form, scores, or rankings.

**BOX 20: EXAMPLES OF QUALITATIVE INDICATORS**

**Overall objective (Impact) level**:

- Degree of citizens' perception of the integrity and trustworthiness of the public service
- Degree of national compliance of labour rights (freedom of association and collective bargaining) based on International Labour Organization textual sources and national legislation, by sex and migrant status
- Level of attainment of national targets established in accordance with Aichi Biodiversity Targets

**Specific objective (Outcome) level**:

- Extent to which policy making on minorities integration by local and national authorities is evidence-based
- Extent to which a comprehensive Public administration reform reporting and monitoring system is in place
- Degree of interoperability of information systems
- Degree of progress in low emission development strategies adoption
- Degree of national's international border management strategy translation into action plans and concrete measures

**Output level**:

- Extent to which new tools and platforms aimed at increasing collaboration and exchange are in place
- Extent to which a mechanism to monitor budget allocation and expenditure is in place
- Degree to which networks of academics and other opinion makers advocate for women's political participation
- Degree to which institutions and partners demonstrate capacity to integrate gender perspectives into policies and measures

*Source: DG NEAR*

#### 9.2.1.2. The intervention logic levels: input, output, outcome and impact indicators

Indicators can be **built at different levels and for different purposes**. Their relevance is very much dependent on their intended use.

Indicators can be categorised according to different elements: **stage of the intervention cycle** (input, activity, output[68], outcome and impact/context), according to the **nature of the indicator** (macro, sector, intervention) and according to the **dimension one wants to look at** (relevance, efficiency, effectiveness, impact, sustainability, coherence, EU added value, etc.).

**BOX 21: INPUT, PROCESS, OUTPUT, OUTCOME AND IMPACT INDICATORS**

***Input indicators*** measure the resources and means provided by donors and implementers.

These indicators typically attract a lot of attention, though they distract programmers and stakeholders from results. Data for input indicators are easy to collect, easy to interpret, and can be timely produced, since they are totally under the beneficiaries' control. Most input indicators measure how fast financial resources are spent, which is what managers in both funding and implementing organizations are really responsible for. However,

[68] The combination of **input**, **process** and **output** indicators can provide evidence of the performance at action level. See DG BUDG, "Evaluating EU Activities: A Practical Guide for Commission Services", 2004 (https://myintracomm.ec.europa.eu/dg/agri/policy/economic_policy_analysis/evaluationandstudies/Internal_Documents/Guide_pub_eval_activities_full_en.pdf).

they say nothing about whether an intervention is achieving its objectives.

Typical examples of input indicators*:*

- *absolute amounts contracted/spent*
- *number of missions having taken place*

***Process indicators***[69] measure what happens during implementation and they mainly focus on the activities execution.

These indicators have weaknesses and have to be used for their purpose, which is to track progress. Some of them (e.g., number of meetings held in a stakeholder consultation) only measure formal elements. Thus, they may signal positive results when only tokenistic compliance has occurred. For example, a managing authority may hold numerous stakeholder consultation meetings—and, yet, still choose not to use the results of the meetings in decision-making. They can relate to production or relations:

I. Production process indicators contribute to address the efficiency issue, if combined with data on outputs at intervention level. They can be aggregated, but require a good data tracking system

Typical examples of production process indicators:

- *Contracting and payment rates,*
- *Rejection rates,*
- *Procurements delays,*
- *Monitoring or audit missions carried out and main findings;*
- *Irregularities and complaints registered and addressed*

II. Relational process indicators may track the political or the implementation processes. These can concern both the operational and the strategic level.

Typical examples of relational process indicators:

- *Number of sub-committee meetings held*
- *Number and quality of consultative platforms held*
- *Number of public/private platform meetings*
- *Number of chapters opened/closed (in the Enlargement context)*

***Output indicators*** show the degree of achievement of the direct products of an activity or set of activities. They are by nature activity-specific. They are typically produced in the framework of the classical project approach, when they follow the implementation of activities.

They are normally easy to report on if implementers collect the necessary data with the necessary level of detail[70].

Output indicators are directly connected with the intervention, unlike outcome or impact indicators, whose value is influenced by other interventions and phenomena[71]. This means that the specific contribution of the

[69] Indicators measuring the substance of processes (e.g., the indicator "quality of discussions in SMSC meetings") are often vague and to be avoided, unless they can easily be transformed into qualitative indicators by defining their meaning and by assigning values. This can be cumbersome and time-consuming and, on occasions, only provide imperfect data on information, which programme managers already have.

Some process indicators, notably the production ones, might be useful only for the parties involved and might be too difficult for stakeholders to interpret.

Data for process indicators are relatively easy to collect, especially at action level, and can be produced during implementation. They can be aggregated in order to monitor performance at strategic level.

[70] For example, every time the output refers to individuals (e.g., trained people), beneficiaries should collect data on gender and on whether individuals belong to minority groups.

supporting donor is clearly identifiable, though their actual realisation and utilisation can depend on other interventions (we can build a premise, but only subject to the conditions that the relevant permits are issued; its use, is subject to the condition that the supplies or the personnel are made available by the partner).

Output indicators may be less informative than outcome indicators on what the intervention is achieving. They are, nonetheless, relevant:

- First, they are within the sphere of control of the stakeholders, which cooperate in implementing the activities.
- Second, output indicators provide valuable information about the presumable quality of the outcomes and about the progress that spurs from the implementation process.
- Third, they can be aggregated.

Typical examples of output indicators*:*

- *Number of people trained;*
- *Number of companies, NGOs, or municipalities supported;*
- *Number of premises built or equipped or maintained;*
- *Number of supplies delivered and installed, events having taken place, laws and regulations drafted.*

***Outcome indicators*** signal whether the short to mid-term desired changes are happening.

They have to be responsive to the intervention, so that they can indicate a change that has taken place also thanks to the intervention. However, they are usually influenced also by other factors: the intervention can build a road, but whether or not it is used and produces benefits on the population and on society depends certainly on its characteristics and on its design, but also on other factors, like cultural habits, system of incentives, or presence of related infrastructures and services. Outcome indicators are key for sector approaches, especially if implemented by means of budget support, when there is no focus on activities but on results to be achieved following the implementation of policy measures. Finding suitable outcome indicators and setting their targets is challenging, since their identification requires reflecting on the strategy set out, the target population, the scope of the intervention, the causal links between inputs (and therefore activities), outputs, and outcomes, the availability of baseline value, the realism of targets set out in terms of value and time—in short, the programme theory.

Typical examples of outcome indicators*:*

- *Time required to enforce a contract*
- *Competition rate on the public procurement market (includes concession award procedures)*
- *Number of households with improved energy consumption classification*
- *Number of fatalities on state and regional roads*
- *Number of pending cases in a Court*
- *Number of legislation entered into force and enforced*

***Impact/Context indicato****rs* signal to which extent the overall (mid and long-term) objectives of a policy or action (the strategic objectives or goals) have been achieved.

Ex ante, these indicators define the macro-economic and sector environment in which we intervene: they act as *context* indicators. Ex post, their evolution suggests whether the environment is moving in the right direction and how it has evolved: they try to show areas where an intervention might have produced an impact.

Impact indicators, however, are not sufficient to attribute the change to the intervention. Only full-fledged ex-

[71] Refer to **FIGURE** 4 : A simplified (linear) depiction of the result's chain of an action

post evaluations may identify to which extent the implemented intervention has actually contributed to a change: changes result from the interplay of many policies, human behaviours, exogenous factors (natural events, main political events, and external shocks).

When referring to the long-term, impact/Context indicators are relevant at Regulation/region/country level, especially for countries that have the perspective of accession, where the combined effect of the political and the financial leverage can be considered such as to influence the relative performance of a country (socio-economic convergence and integration, political-institutional transformation). In the mid (intermediate) term, they are also relevant at intervention level.

Typical examples of impact indicators*:*

*Long-term:*

- *Real Gross Domestic Product growth*
- *Public debt/GDP (% )*
- *Unemployment rate (%)*
- *Foreign Direct Investment per capita*
- *Composite indicators such as the Inequality-Adjusted Human Development Index*

*Intermediate term:*

- *Number of trade barriers removed*
- *Number of people being held without trial*
- *Degree of compliance with intellectual property legislation*
- *Composite indicators such as the World Bank "Distance to Frontier" index*

*Source: DG NEAR*

It is important to highlight that **these distinctions do not relate to the indicator per se**. **The same indicator may measure different phenomena in different circumstances**. For example, timely expenditure of planned resources functions as a process indicator in most cases—and tells next to nothing about whether the intervention is producing positive or negative results. It acquires a different meaning when used in connection with an activity aiming at increasing the capacity of partner country's implementing agencies. In this case, the rate of expenditure in the entire intervention (or in sectorial interventions) may measure whether implementers have learned the procedures required by the European Commission. It, therefore, might be used among outcome indicators of increases in implementation capacity.

### *9.2.2. How to define an indicator? What are its components?*

An indicator comprises[72]:

- a **title**: the name of the indicator, e.g., length of rehabilitated road
- a **definition**, which describes the unit of measure, what is observed and how the value of the indicator is calculated at (e.g., km of rehabilitated county road/total km of county road in the area where the intervention operates)
- a **baseline**: the reference value at the start of the intervention (or the latest available) against which progress will be assessed
- **a milestone[73]**: the path towards the final target. They can be more than one if the intervention lasts for a long time

[72] The table in Annex 6 synthesises the information which should be provided for each indicator**. It is unlikely that beneficiaries/partners are able to provide all this information**: many fields will probably require many re-iterations to be filled. Nevertheless, it is useful to at least discuss these points in identifying the indicators, in validating them, and in reviewing them. This allows to at least identifying points where the entire process may encounter problems before the problems arise.

[73] At the beginning of the current programming period the milestone year for IPA II was set at 2017.

- a final **target**, which signals how much change is expected and in what direction
- The **reference period** (usually year) to which the baseline, milestone(s) and final target values refer
- a **source** of verification
- indication of the **frequency** by which the indicator is or needs to be updated
- and the **beneficiaries' responsibilities** for collecting the data and building the indicator.

**TABLE 6: EXAMPLE OF INDICATOR COMPONENTS**

| Title | Definition | Baseline | | Milestone | | Target | | Sources of information | Data collection responsibilities & frequency |
|---|---|---|---|---|---|---|---|---|---|
| | | Value | Year | Value | Year | Value | Year | | |
| % of women in senior positions of trade unions | This indicator is the percentage of women in senior management positions (public sector; at least head of unit post). | 15% | 2010 | 20% | 2015 | 25% | 2020 | Department of Statistics of the given country | Implementing partner, yearly |

*Source: DG NEAR*

When relevant, indicators need to be **disaggregated according to different dimensions**: gender, age, a territorial level to which it refers (e.g., country, region, province), income quintiles/deciles, etc. When an indicator is fit for disaggregation, it is important **to think about different variables to decide** on whether or not to finally consider this given indicator: the disaggregation cost (resources available), data quality and the data collection process, the frequency of reporting, etc.

**BOX 22: GENDER EQUALITY DIMENSION IN INDICATORS**

European Union policy mandates that **gender equality be embedded in thinking about indicators**, exactly as gender is embedded in our everyday social experience.[74] It should not be an afterthought when designing an indicator system. After all, the world is gendered, so any representation of changes in relevant dimensions of the world should include gender equality.

There are two ways to include gender equality into an indicator system: first, by including **gender thinking** in all the steps involved in identifying indicators and, second, by including **gendered indicators[75]& [76]**. There are three types of gendered indicators: **sex disaggregated indicators**, **ad hoc indicators**, and **indirect indicators**.

**Sex disaggregated indicators**: indicators measuring the number of individuals which have been involved in or been impacted by the intervention should be sex disaggregated, i.e., should separately report the number (or

[74] The SWD (2015) 182 final - "Gender Equality and Women's Empowerment: Transforming the Lives of Girls and Women through EU External Relations 2016-2020" provides a framework for action and measures for EU actors to deliver results and to report transparently on progress and setbacks in relation to the following objectives:

- Ensuring girls' and women's physical and psychological integrity
- Promoting the economic and social rights / empowerment of girls and women
- Strengthening girls' and women's voice and participation
- Shifting the Commission services' and the EEAS' institutional culture to more effectively deliver on EU commitments.

[75] A non-exhaustive list of indicators that can be used to measure contextual progress, and/or EU contribution to change is provided in Annex 1 of European Commission, 2015d.

[76] See examples of gender related impact indicators in Annex 8 and of gender related outputs and outcome indicators in Annex 9.

percentage) of men and women, or the value of the indicator for men and women separately (such as unemployment rate).

*Examples are indicators measuring the number of trained people, education achievement indicators, or employment indicators (both those measuring the employment generated by the action and those measuring employment in the general population). Another type of sex disaggregation may refer to issues such as Regulatory simplification. For example, it is possible to identify how long it takes to enforce a contract or to start a company for women in a given context.*

Unlike statistical indicators of employment, action-level indicators present difficulties because they are based on data from monitoring or administrative systems. Administrative forms may not include the information on the applicant's gender or the trainers' gender.[77] This lack of information might be difficult to fix because of privacy norms or of people objecting to declaring their gender.

**Ad hoc indicators**: some interventions may try to increase the weight of women in decision-making roles or the percentage of women-owned enterprises or the friendliness of some services to women (e.g., by staffing anti-domestic violence centres with women police officers). Indicators, therefore, directly measure these dimensions.

**Indirect indicators**: such as coverage of residential care or of childcare, do not openly refer to women or men, but measure dimensions, which in turn affect gender equality. It is necessary to have a sufficiently developed theory (see part 1) to support the claim that the indicator is linked to gender equality.

*Source: DG NEAR*

**Indicators must be expressed in neutral terms**: **they signal a change, not the direction of the change** (as objectives do). For example, if there is an objective to decrease the number of gender based violence cases, the corresponding indicator might be 'Number of gender-based violence cases reported', and not 'Decreased number of gender-based violence cases reported'.

### *9.2.3. What are the characteristics of indicators?*

"Good" or "bad" indicators do not exist in nature. **The quality of an indicator depends on the purpose it serves, on the nature of the intervention, on the intervention stage, and on the development and maturity of the performance system**[78].

When defining indicators, there are no unique or universal principles to be followed. The Better Regulation establishes that, when relevant, all indicators should be RACER.

**TABLE 7: RACER PRINCIPLES FOR INDICATOR'S DEFINITION, AS PER THE BETTER REGULATION**[79]

| Principle | Definition |
|---|---|
| **Relevant** | Closely linked to the objectives to be reached. They should not be overambitious and should measure the right thing |
| **Accepted** by staff, stakeholders | The role and responsibilities for the indicator need to be well defined |

[77] Trainers' gender is relevant when providing training on sensitive issues, e.g., when training police or judiciary professionals on how to deal with victims of gender-based violence or when trying to contrast discrimination by providing role-models, e.g., when contrasting gender segregation in professions or at upper echelons of decision making.

[78] Even a "bad" indicator may be a start in a **learning process**, since what is not detectable during a planning/programming period may become measurable over time, as interventions improve collection of administrative and statistical data in partner countries and experience, studies, and evaluations clarify the concepts of interest. For example, an effective performance system cannot rely solely or even mainly on input and process indicators, because they fail to provide information about what the intervention is achieving. However, at the very beginning of the efforts, they might be all that is feasible to define and to report against, and allow the partner countries to start gaining experience in monitoring through these indicators.

[79] Definitions have been further elaborated.

| Principle | Definition |
|---|---|
| **Credible** for non-experts | Unambiguous and easy to interpret. Indicators should be simple and robust as possible. If necessary, composite indicators might need to be used instead – such as country ratings, well-being indicators, but also ratings of financial institutions and instruments. These often consist of aggregated data using predetermined fixed weight values. As they may be difficult to interpret, they should be used to assess broad context only. |
| **Easy** | To monitor (e.g. data collection should be possible at low cost).<br>Built, as far as practicable, on available underlying data, their measurement not imposing too large a burden on beneficiaries, on enterprises, nor on the citizens |
| **Robust** | Against manipulation (e.g. administrative burden: If the target is to reduce administrative burdens to businesses, the burdens might not be reduced, but just shifted from businesses to public administration).<br>Reliable, statistically and analytically validated, and, as far as practicable, complying with internationally recognised standards and methodologies; |

*Source: Better Regulation*

The Better Regulation also acknowledges that if necessary, **proxy indicators** might need to be used to represent a phenomenon in the absence of a direct indicator. In these cases, it recalls the necessity to correctly understand the underlying causal links and the limitations of using proxies.

Other than RACER, other complementary principles are widely used at the international level; this is why they are provided here.

**TABLE 8: SMART PRINCIPLES FOR INDICATOR'S DEFINITION**

| Principle | Definition |
|---|---|
| **Specific** | Target a specific area for improvement |
| **Measurable** | Quantify or at least suggest an indicator of progress. |
| **Attainable/achievable** | Make sure that it can be reached |
| **Realistic** | State what results can realistically be achieved, given available resources |
| **Time-bound** | Specify when the result(s) can be achieved. |

*Source: https://en.wikipedia.org/wiki/SMART_criteria*

**TABLE 9: ADDITIONAL PRINCIPLES FOR INDICATOR'S DEFINITION**

| Principle | Definition |
|---|---|
| **Reasonable** | Capturing the essence of an outcome according to a reasonable argument about which features of the outcome they can and cannot represent; |
| **Normative** | Having a clear and accepted normative interpretation (i.e. there must be agreement that a movement in a particular direction or within a certain range is a favourable or an unfavourable result) |
| **Debatable** | Timely and openly available to a wide public, with room being built for public debate and for their own revision when needed and motivated. |

*Source: "Outcome indicators and targets" produced for DG Regional Policy by an expert group led by F. Barca and P. McCann, June 2011*

These sets of principles refer to each indicator. It is important, however, to consider indicators within the indicator system. This implies some considerations, which are specific to the indicator system:

- Strike the right balance between including **several specific indicators** and selecting a small number of **common indicators** in order to aggregate results data and cross-**compare** the **performances and report**. This requires that the way the data are collected and processed is the same across interventions and countries—which might need harmonising and negotiating indicators with the various partner countries. Comparison and aggregation serve different purposes and interest different stakeholders:
    - *Aggregation*: Funding agencies and national authorities may need to aggregate data in order to demonstrate performance at national, regional (e.g., the Western Balkans, ENI East, ENI South), or instrument (e.g., ENI/IPA II) level. These aggregations may possess limited meaning and interest for beneficiaries.
    - *Comparison*: all stakeholders benefit from comparing performance across interventions, areas, implementing agencies, and partner countries. Even implementers and beneficiaries can make good use of information about their relative performance.
- Ensure that the data are used and that the demand for information from the indicator system increases over time. As for all knowledge, **utilisation of data is important**.
- Ensure that there is explicit support and that there are sufficient resources since the beginning and especially during intervention implementation. In each planning/programming period, **attention should focus more on operating the indicator system than on building it**.
- **Do not identify too many indicators** in order to keep the cognitive burden on all stakeholders within acceptable limits. Using data from indicator systems in programme management requires cultural changes among all decision levels at EU, national, and local levels. "It has been shown that in a situation of decision-making, a person cannot take into account more than about ten indicators at once. When there are too many indicators decision-makers are swamped with an excess of information." (European Commission, Evalsed: 87). **Less is better.**

### *9.2.4. Indicator frameworks for ENI & IPA II*

In response to a greater focus on results being impressed across European Union institutions, and at international level, the Commission has placed a major effort on disclosing results and measuring performance.

In the case of ENI, the framework is provided by the EU International Cooperation and Development Results Framework (the EU Results Framework) launched by the Commission's Staff Working Document of 26 March 2015[80]. It aims at strengthening the capacity of the relevant DGs (DEVCO and NEAR) to monitor and report progress made in achieving results, thus enhancing accountability, transparency and visibility of EU aid[81]. It was established with clear, pre-defined and where appropriate, country-specific and measurable indicators.

In the case of IPA II the reference is the IPA II performance framework. It aims at providing a coherent context for performance measurement and for improvement of relevant monitoring and reporting processes and tools for pre-accession assistance. Progress towards achievement of the objectives is measured, monitored and assessed by means of pre-defined, clear, transparent and, where appropriate, country-specific and measurable performance indicators.

[80] European Commission, 2015d.

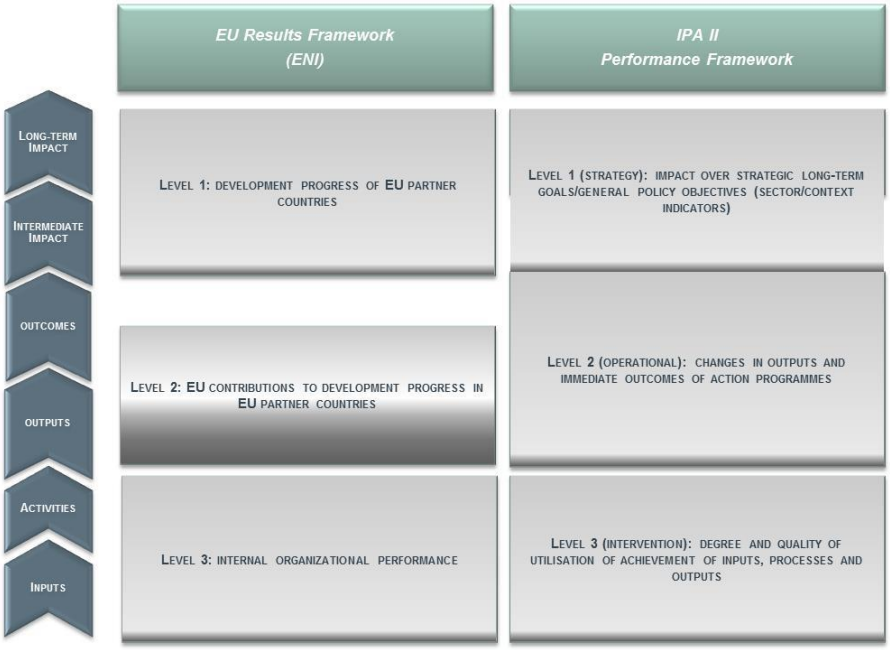
[81] The first report based on the new "EU International Cooperation and Development Results Framework" was published in July 2016 and can be accessed at: https://ec.europa.eu/europeaid/news-and-events/eu-publishes-key-results-its-international-cooperation-and-development-activities_en.

Each of these frameworks aims at aggregating results. They include specific quantitative and qualitative indicators, grouped into three different levels that are linked to a result's chain. It is possible to identify similarities and equivalences (refer to **FIGURE 10: INDICATOR FRAMEWORKS FOR ENI AND IPA II**).

- The IPA II and ENI level 1/strategic level indicators are reflected in the Indicative Strategy Papers (ISP) (for IPA II) and in the Single Support Framework (SSF) or equivalent document (for ENI). They measure the attainment of overall objectives in terms of impacts. Only for IPA II, a common set of strategic indicators has been proposed by DG NEAR for the ISPs. Beneficiaries have been asked to set targets for 2020 with reference to them.
- Level 2/operational level indicators measure the achievement of the specific objectives in the each action programme, which hierarchically contribute to the achievement of level 1 indicator. No list of specific indicators has been proposed in this case, the indicators have to reflect the specificity of the interventions. Nonetheless, for reporting purposes, a list of common indicators (mostly outputs) has been drawn under the EU results Framework for ENI countries.

**The indicators included in these frameworks are to be used in corporate reporting. This means that when relevant, they are expected to be included in actions' planning/programming exercises**. Their use if for the time being not mandatory but for the Commission to be able to report it is highly suggested to use them.

**FIGURE 10: INDICATOR FRAMEWORKS FOR ENI AND IPA II**




*Source: DG NEAR*

#### 9.2.4.1. ENI

As mentioned, ENI actions utilise the categories defined in the EU result's Framework defined by DG DEVCO[82]:

**Level 1 indicators** track development progress/ strategic indicators, concerning essentially mid and long term impacts/context at both macro and sector level;

**Level 2 indicators** signal the changes intervening in outputs (mostly) and outcomes (to the possible extent) identifying those aspects on which the EC is contributing to results by means of its financial support; and;

[82] Refer to *Launching the EU International Cooperation and Development Results Framework*, SWD(2015) 80 final.

**Level 3 indicators** record action/organisational performance, focusing on resources and activities used and the way the financial cooperation is deployed (quality of planning/programming, use of agency Human resources, speed of implementation).

#### 9.2.4.2. IPA II

Regarding IPA II, DG NEAR undertakes the approach of using performance indicators as part of the performance framework. It means that a selected number of indicators pertaining to different categories (strategic, operational and action levels) are to be monitored to track evolution against the targets set out.

##### What are the specificities in terms of IPA II indicators?

For IPA II actions, there are three levels of indicators (see **FIGURE 11: IPA II INDICATORS TYPOLOGY AND WHERE TO USE THEM**, synthesising indicator types and their use in documents):

**Strategic indicators (level 1)** provide information about the change in the dimensions, which are connected with long-term goals/general policy objectives. They track the level of achievement of actions' contribution to **impacts**—of course, they cannot directly measure it, because as mentioned under **FIGURE 4 : A SIMPLIFIED (LINEAR) DEPICTION OF THE RESULT'S CHAIN OF AN ACTION**, impacts are to be considered as **the intermediate to long-term changes** in the political, social, economic and environmental global context which will stem from interventions of all relevant actors and stakeholders and **which will be indirectly influenced by IPA II[83].**

Strategic indicators can be at *macro* level or at *sector* level:

- **Macro (context/impact) indicators** synthesise changes in the business environment, the legislative framework, the institutional setting, and the social performance. Indicators at macro level are to be used in **annual progress reports** assessing the overall evolution in a beneficiary country. As all indicators, there are both quantitative and qualitative macro indicators. Macro indicators, which can be a mix of impact/context, are suitable for the **Programme Statement**, at an aggregated level; some of them also for **(Multi-) Country Strategy Papers**.
- **Sector indicators** capture the progress in implementing a broader reform agenda that has been identified as a priority in the relevant (Multi-) Country Strategy Paper. Performance can then be tracked by means of sector indicators, looking at the improved performance of the sector, the results that can be felt at the level of the population and not only the direct beneficiaries of the actions. They are associated to financial assistance, which targets specific sectors and/or is provided on the basis of programmes addressing the needs in a given policy area or sector. Some key indicators can be reflected in the **(Multi) Country Strategy Papers**.

**Operational indicators (level 2)** capture changes in **outputs** and **outcomes**. The connection between the action and these indicators is closer than for strategy indicators, but at the outcome level, the mere change in indicators does not suffice by itself in showing that the change was caused by the action (since other interventions also influence outcomes; that is a job for full-fledged evaluations). Nevertheless, they provide useful knowledge about whether the situation is evolving in the right direction. Coupled with other information, this knowledge can greatly improve the ability of implementers and funders to steer the action. Operational indicators are basically output indicators with some process and outcome indicators, and should be formulated in documents for action design, such as **Action programmes/documents[84]**.

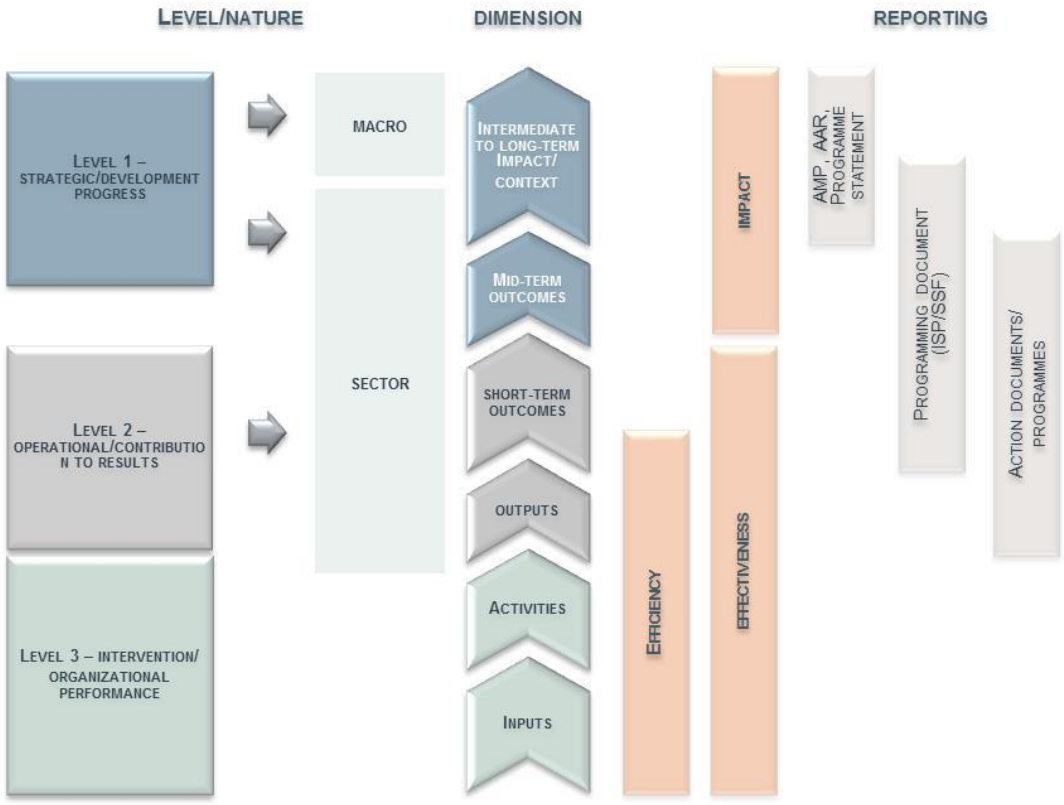
[83] The best example is provided by pre-accession assistance: its overall impact can only be reasonably assessed in connection with the whole policy framework. This is because the accession perspective represents a powerful leverage of change, while financial assistance only aims at facilitating the socio-economic, cultural and institutional development which is necessary in order to bring about the desired political change, i.e. preparation for accession.

[84] Whilst in the EU Results Framework developed by DEVCO Level II indicators are mostly focused on output indicators, Operational indicators proposed in relation to the performance framework elaborated for IPA II try to address more the outcome dimension.

**Intervention indicators (level 3)** reflect the extent to which **inputs** are used to produce activities, which in turn are expected to produce outputs in a certain quantity, quality and time; they are situated at action level but also at management and organisational in the case of political inputs for example, that are also necessary to implement your action. They consist of input and process indicators. These indicators reflect the efficiency of the organization and the way by which resources are deployed and implementing mechanisms and modalities) take place.

The whole set of indicators (strategic, operational and action) constitutes the **performance/results framework**.

**FIGURE 11: IPA II INDICATORS TYPOLOGY AND WHERE TO USE THEM**




*Source: DG NEAR*

### 9.3. Who should take part in the process?

Building and maintaining an indicator system is a collective endeavour. It requests active participation from programme managers (both from the HQ and the EUDs), sector experts (HQ CoTEs, EUDs, operational structures from partner countries), representatives of the national statistical offices, etc.

There are also other stakeholders, who are directly/indirectly affected by the indicators:

- Representatives of the CSOs in the partner countries, business associations, academic institutions, and final beneficiaries of ENI/IPA II
- General Secretariats of Government/ Deputy Prime Minister´s or Prime Minister´s Offices, Ministries of Finance, European Integration, EU Affairs, Development and Cooperation in ENI/IPA II countries etc.;
- EU Institutions - EU Member States; EU Parliament; EU Council; European Court of Auditors.
- IFIs

- International organisations, etc.

The specific responsibilities on defining, monitoring, reporting on performance indicators are presented in the following tables (one for ENI and one for IPA II). The tables illustrate the roles of the different actors in the main steps/activities in the planning/programming, implementation, monitoring, reporting process. It is not aimed at providing a comprehensive overview of the full set of responsibilities, but rather to present the collective nature of the process and the necessary participation from the relevant stakeholders. The table is, therefore, indicative.

**TABLE 10: ROLES AND RESPONSIBILITIES OF ENI ACTORS ON INDICATORS PLANNING/PROGRAMMING, IMPLEMENTATION, MONITORING AND REPORTING (STATE OF PLAY OF THE CURRENT SITUATION IN 2016)**

| Actor | Activities in developing, collecting data, monitoring, reporting indicators | | | | |
|---|---|---|---|---|---|
| | Indicative strategy papers indicators (Single Support Framework) indicators (level 1) | Action programme logical framework identification and formulation of indicators (baseline, benchmarks, targets, source of information, etc.) | Data collection system informing the indicators | Monitoring progress of the indicators | Reporting on indicators |
| **EC & EUDs** | | | | | |
| DG NEAR HQ Evaluation Service | Quality control | Quality control | | Quality control (upon request) | Quality control (upon request) |
| DG NEAR HQ Country/Regional programmes managers/ DG NEAR HQ (CoTEs) | Prepare | Quality control | NA | | |
| EUDs | Prepare | Prepare | Compile data with support of the ROM end of project results reporting experts | | Reports on some level 3 indicators (i.e. EAMR) |
| DG DEVCO | | NA | Prepares (level 1 indicators) | Quality control (ENI level 2 indicators) | Aggregate and Report (level 1, 2 and 3 indicators) against the EU corporate Results Framework indicators and the specific country level 2 indicators |
| **National authorities** | | | | | |
| ENI Aid Coordinators | Consulted | Contributes | | | |
| Sector national institution[85]/Operational structures from partner countries[86] | Consulted | Contributes | Collects | | Informed |

[85] The relevant and the roles of this institution depend on the degree of decentralization in the management of financial cooperation.

[86] In some cases, this might be the sector national institution.

| Actor | Activities in developing, collecting data, monitoring, reporting indicators | | | | |
|---|---|---|---|---|---|
| | Indicative strategy papers indicators (Single Support Framework) indicators (level 1) | Action programme logical framework identification and formulation of indicators (baseline, benchmarks, targets, source of information, etc.) | Data collection system informing the indicators | Monitoring progress of the indicators | Reporting on indicators |
| Representatives of the national statistical offices | Consulted | Contributes | Collects | | |
| ENI Committees[87] | Provide opinion | Provides opinion | NA | NA | Informed |
| EC College | Approves | Approves | | | Informed |

*Source: DG NEAR*

**Legend:**

- Aggregate – consolidation of information;
- Approve – adoption of the documents by the competent authorities;
- Check – performing consistency checks for the coherence and complementarily of the indicators;
- Collect – the process of gathering data for monitoring and reporting on the indicators;
- Consult – consultation with the key actors (line ministries, international organisations, etc.) on the national political priorities;
- Contribute – contribution to the identification of indicators and related information, ensuring the credibility of the data, methodologies,
- Data collection system – activities, referring to collection of data for feeding the analysis of the indicators
- Informed – dissemination of information on the adopted/reported indicators;
- Monitor – overviewing, monitoring the progress in achieving the indicators;
- Monitoring progress of the indicators – activities, referring to tracking the progress in achieving the results, measured by indicators
- Prepare – drafting planning/programming and planning documents and identifying the relevant indicators;
- Quality control –control/check of data quality in view of ensuring consistency and reliability of results and its coherence with the interventions' objectives.
- Report – reporting on the progress of the indicators, using relevant sources;
- Reporting on indicators – activities, referring to reporting on indicators in reports at different level (EUD reports, DG NEAR HQs reports)

[87] The Committees are those mentioned in the Regulation No 232/2014 establishing an European Neighbourhood Instrument.

**TABLE 11: ROLES AND RESPONSIBILITIES OF IPA II ACTORS ON INDICATORS IDENTIFICATION, PRODUCTION, MONITORING AND REPORTING (STATE OF PLAY OF THE CURRENT SITUATION IN 2016)**

| Actor | Activities in developing, collecting data, monitoring, reporting indicators | | | | | |
|---|---|---|---|---|---|---|
| | **Indicative strategy papers indicators (level 1) Planning/programming documents** | **Sector planning documents indicators (level 2)** | **Action programme logical framework identification and formulation of indicators (baseline, milestones, targets, source of information, etc.)** | **Data collection system informing the indicators** | **Monitoring progress of the indicators** | **Reporting on indicators** |
| **EC & EUDs** | | | | | | |
| DG NEAR HQ Evaluation Unit | Quality control | Quality control | Quality control | | Quality control (upon request) | Quality control (upon request) |
| DG NEAR HQ | Prepare (Directorate A country units) | Quality control (country units/CoTEs, A3), A4) | Quality control | Prepares (level 1 indicators – A3) | Gathers and assesses IPA level 1 indicators Quality control (all other IPA levels indicators) | Aggregates and Reports (A1/A4) |
| DG NEAR HQ Country/Regional programme managers/DG NEAR HQ (CoTEs) | Prepare | Quality control | Quality control | NA | | |
| EUDs programme manager | | Quality control | - Quality controls<br>- Encodes the approved actions indicators in MIS | - Collects output indicators/Aggregates (in direct management) and Reviews (in all management modes) | - Quality control (level 2 and 3 indicators) in indirect management<br>- Implements (level 2 and 3) for direct management | Reports |
| **National authorities** | | | | | | |
| NIPAC representative | Consulted | Quality controls | Prepares | Collects output (in indirect management), and outcome indicators Aggregates and Quality control (in indirect management) | Implements (level 2 and 3 indicators) | Report (level 2) |
| Sector national institution[88]/Operational | Consulted | Prepare | Prepares | Collects | Implement (level 2 and 3 indicators) | Informed |

[88] The relevant and the roles of this institution depend on the degree of decentralization in the management of financial cooperation.

| Actor | Activities in developing, collecting data, monitoring, reporting indicators | | | | | |
|---|---|---|---|---|---|---|
| | Indicative strategy papers indicators (level 1) Planning/programming documents | Sector planning documents indicators (level 2) | Action programme logical framework identification and formulation of indicators (baseline, milestones, targets, source of information, etc.) | Data collection system informing the indicators | Monitoring progress of the indicators | Reporting on indicators |
| structures from partner countries[89] | | | | | | |
| Representatives of the national statistical offices | Consulted | Contributes | Contributes (IPA and ENI) | Collects (IPA II and ENI) | | |
| IPA II Committees [90] | Provide opinion | NA | Provides opinion | NA | NA | Informed |
| EC College | Approves | | Approves | | | Informed |
| IPA II Monitoring Committee | | | | | Monitors (level 2 and 3 indicators) | Reviews (level 2 and 3 indicators) |
| Sector Monitoring Committees (IPA II)/ Joint Monitoring Committees (IPA II CBC) | | | | | Monitors (level 2 and 3 of indicators) | Reviews (level 2 and 3 of indicators) |

*Source: DG NEAR*

**Legend:**

- Aggregate – consolidation of information;
- Approve – adoption of the documents by the competent authorities;
- Check – performing consistency checks for the coherence and complementarily of the indicators;
- Collect – the process of gathering data for monitoring and reporting on the indicators;
- Consult – consultation with the key actors (line ministries, international organisations, etc.) on the national political priorities;
- Contribute – contribution to the identification of indicators and related information, ensuring the credibility of the data, methodologies,
- Data collection system – activities, referring to collection of data for feeding the analysis of the indicators
- Gathers and assesses –monitoring and reporting on the progress of the indicators, using different reports;
- Informed – dissemination of information on the adopted/reported indicators;
- Level 1 – In the context of IPA II, level 1 indicators are strategic (impact/context) ones.
- Level 2 - In the context of IPA II, level 2 indicators are operational (outcome and output) ones.
- Level 3 - In the context of IPA II, level 3 are organisational indicators (process indicators and input when relevant).

[89] In some cases, this might be the sector national institution.

[90] The Committees are those mentioned in the Regulation No 231/2014 establishing an IPA II.

- Monitor – overviewing, monitoring the progress in achieving the indicators;
- Monitoring progress of the indicators – activities, referring to tracking the progress in achieving the results, measured by indicators
- Prepare – drafting planning/programming and planning documents and identifying the relevant indicators;
- Quality control –control/check of data quality in view of ensuring consistency and reliability of results and its coherence with the interventions' objectives.
- Report – reporting on the progress of the indicators, using relevant sources;
- Reporting on indicators – activities, referring to reporting on indicators in reports at different level (NIPACs report, EUD reports, DG NEAR HQs reports)

### 9.4. Practical steps in defining, verifying, using indicators for reporting for Action programmes

The development of effective performance indicators spans more than a planning/programming period: development of effective sector-level M&E systems, starting with sector-level performance indicators, is a long-term iterative process. Time is necessary to change internal processes and inter-institutional relations, especially, to learn from experience and to advance knowledge in measuring quantitative indicators and in using qualitative ones. Gradually, national statistical systems improve and administrative data collection advances. This brings measurement even in areas in which measurement is not possible during one planning/programming period. Conceptual advances and refinement of interventions lead to the development of better and trust worthier qualitative indicators. It is important to keep track of the learning that occurs during one planning/programming period, because the clarification and testing of programme theories, and the establishment and operation of functional data collecting systems (for both statistical and administrative data) take a long time and produce their effects over the long term.[91]

This section deals with the practical issues involved in identifying **indicators** to monitor and evaluate interventions. The work to be done requires in-depth thinking about the concrete activities that will be implemented and the affected individuals' and organisations' reactions. As mentioned under 8.3.2, the key moment to define indicators is during programming, when the intervention logic is devised. **Having said this, identification and review of indicators may occur as well during monitoring and evaluation. Indeed, if indicators defined during planning/programming are expected to be used in monitoring and evaluation exercises, depending on the purpose and scope of these exercises, additional indicators might be defined and existing indicators amended/revised.**

The section describes the **main steps** involved in defining indicators, from reconstructing the intervention logic to designing the data collection system. As the work progresses from one step to the other, the information must be recorded and fed in fiches (one per each indicator, according to the example in Annex 6).

**FIGURE 12: IDENTIFYING INDICATORS: MAIN STEPS**


The most important tasks in building indicators are those which immediately precede and follow the identification and definition of the indicators and the construction of the performance system: **assessing the merits and pitfalls of existing indicators** and, after having identified new ones, **operating the indicator system**, **adjusting it** as difficulties arise and new opportunities unfold, and **using the data** coming from the system.

#### *9.4.1. Step 1 – Reconstruct the intervention logic*


[91] "First, are indicators actually reported on and do they prove to be technically sound? Second, are these indicators meaningful in the decision making process, are the useful for guiding resource allocation or policy making? In the long term, growing M&E experience might lead countries to upgrade to more outcome-oriented indicators or develop more sophisticated and robust M&E systems. Ultimately, countries would ingrain the practice of linking monitoring and evaluation to support evidence based policy making" (The World Bank, 2014: 35).

Identifying indicators is usually prescribed as an integral part of developing an intervention. Performance indicators need to be identified, designed and defined at the planning/programming stage, though they can be subsequently updated in view of the experience and of data availability.

Real practice, however, rarely fits in with these prescriptions. Sometimes, identifying indicators occurs as an **iterative process**, with indicators being set up at each step of planning/programming, from Country Strategies and Sector Strategies to Interventions, usually after the main lines of the strategy have been decided. Often, **only after interventions have been conceived, the moment comes when identifying the log frame and the indicators becomes an urgent imperative**.

**This real life sequence, though it is neither optimal nor desired, provides a very good opportunity to improve interventions** by reconstructing the intervention logic, identifying the need for improvements in the intervention, and agreeing on operational courses of actions with implementers. It also allows for identifying what really is important in the intervention, and to apply ingenuity to monitoring and measuring it. Since this further iteration happens at a time when a clearer view of the challenges ahead is available, it results in a **more robust intervention logic**.

As mentioned in the previous chapter, planning/programming, monitoring and evaluation are closely interlinked. Once the intervention logic is defined, the intervention's internal monitoring system needs to be conceived as well[92].

Reviewing the intervention logic starts from the **logical framework**—which provides a snapshot of intervention activities, outputs, outcomes and impacts and clarifies the assumptions made. As seen in 8.1.1, **intervention logics** go one step forward (and many levels in depth): they open up the **black box** of interventions, and fill in the gap between each activity and the objectives, describing the **causal mechanisms linking inputs, activities, outputs, outcomes and impacts**. This is the **first step in identifying indicators**: through presenting and discussing the logic behind the intervention, one has to lay out exactly what it is being measured and what the major assumptions are concerning the contribution of the intervention. Therefore, they also help identify the links on which knowledge is needed.

### *9.4.2. Step 2 – Identification/selection of indicators*


*Description*

Once the relevant results' statements (e.g., improvements in public sector capacity) have been defined, it is necessary to find out **which observable variable captures the essence and the direction of change**. This is done via the identification of indicators. Too many indicators make reporting onerous and overburden decision-makers. Therefore, **frugality** must be the norm, resulting in a limited number of indicators, which impose relatively little burden to be repeatedly quantified during and after the life of the intervention.

As mentioned under 9.3, **the identification of indicators is a collective endeavour**. It should include representatives of the different stakeholders (EC/EUD, partner country, etc.), which during implementation have the responsibility of collecting, analysing and reporting on data. This is a delicate task, since individuals and organisations adapt to the way information is collected about their activities, modifying their

[92] Refer to § **9.4.7 STEP 7 – DESIGN THE DATA COLLECTION SYSTEMS**.

behaviour[93]. The task of identifying the right indicators and targets is specifically delicate in the framework of budget support operations, where their actual availability, relevance and attainability are critical, conditioning the disbursement of tranches.

**The identification/reconstruction of the intervention logic shows its utility here**: it helps **thinking in concrete terms**, about what is supposed to happen and who acts, rather than about what the intervention finances.

It is necessary, at this point, to identify which indicators measure i) the production of the outputs that are crucial for the causal chain to progress and for the mechanisms to be activated and the contribution to ii) outcomes and iii) impacts. The actors identifying the indicators should decide **which dimensions should constitute the focus of attention.** This implies delicate decisions, because the very fact of selecting and following the evolution of one dimension of an intervention attracts attention on it—to the detriment of other dimensions.

In this framework, it is crucial for outcome and impact indicators to choose areas in which **data is collected by the partner country's public authorities responsible for implementation** (for administrative data) **or by a credible organisation** that routinely produces the data and provides it in an open format (e.g., a national statistical office). If there is no such data collection and processing in place, then **the intervention should make specific provisions** to ensure it is started on a permanent basis—rather than including ad hoc surveys or studies which might be of limited use.

The **principles for indicator's quality** (refer to § **9.2.3 WHAT ARE THE CHARACTERISTICS OF INDICATORS?**) show their operational utility in choosing indicators: they provide the basic criteria.

**Indicators defined during planning/programming might be revised** if during the implementation of the intervention it is found out that:

- there are finally no reliable and sound data sources for providing a value of measurement for one or more indicators;
- one or more indicators are finally not that relevant for grasping the intervention performance;
- new indicators are needed.

**Indicators defined during programming are to be used to the maximum extent possible in monitoring and evaluation exercises.**

Finally, **indicators should be set to the possible extent on the basis of the indicative list of ENI & IPA II performance indicators** (included in Annexes 8 to 10). Indeed, these are indicators that are only suggested for consideration. They are not compulsory (with the exception of the strategic indicators included in the Indicative Strategy Papers under IPA II), and they do not pretend to represent an exhaustive list. Beneficiaries and/or programme managers are only expected to get inspired by them, whenever relevant. Nonetheless, for corporate reporting purposes, the Management might decide to select a set of these indicators, like it is already the case for those included in the EU Results Framework covering the Neighbourhood countries.

When indicators are reflected in a performance measurement table annexed to the Action Document, the indicators should be aligned with those of Logical Framework Matrix (Log frame).

[93] For example, school teachers may focus their efforts on the best students, the worst performing ones, or the average students depending on how class performance is measured. This adaptation may be positive or negative depending on how the indicators are chosen: for example, introducing gendered indicators (e.g., sex disaggregation of students' performance) may have positive effects by cross-comparing incentives on teaching boys' and girls' performances. **Past experience** is key on this issue: it will be necessary to assess which distortions took place—evaluations may address this issue.

**BOX 23: EXAMPLES OF SUPPORT QUESTIONS AND TIPS IN IDENTIFYING INDICATORS**

The key **questions** to be asked when choosing indicators are developed on the basis of the principles for quality of indicators (Annex 7).

**TIPS**

When defining the indicator, bear in mind that **they should be formulated as measurement** (e.g. government effectiveness (rank), level of effectiveness of PAR strategies, etc.), **not result** (e.g. PAR strategies/actions plans designed in a way that they can be implemented) **or objective** (e.g. increase of capacities to elaborate and implement government budget), etc.

Even slight differences in definition can have important implications both in terms of accuracy/reliability and data collection costs. **A clear definition, including the unit of measurement, is particularly important as data (pertaining to specific indicators defined in the EU result framework and in the IPA II Performance framework; refer to 9.2.4 INDICATOR FRAMEWORKS FOR ENI & IPA II) need to be aggregated at corporate level.** Aggregation of data can also become problematic if the definition is not provided or if indicators are not calculated on a consistent basis.[94]

**Do not define too many indicators for a single result**. Strategic aspects, as well as available resources for data gathering and analysis need to be considered when establishing the list of indicators. Of course, the final selection needs to ensure that the scope of indicators do actually allow measuring progress or achievement on an intended result.

In some cases, where it is difficult to find one/two indicators to capture the given result, it is **advisable to use composite indicators**. They shall include several individual ones.

When defining indicators, **do consider data collection responsibilities.** This may have an influence on the final selected list.

**Look for a balance of quantitative and qualitative information.** Whether quantitative or qualitative indicators are appropriate (and feasible to collect) will depend on the nature of the intervention's objectives, as well as other contextual factors. Quantitative measures of achievement can have particular advantages (ease of specification, aggregation, and comparison), however they may also give an unrealistic sense of precision and more importantly, an incomplete vision of actual achievements. Indeed, **in particular in the case of DG NEAR intervention's portfolio, the political and institutional nature of support does not fit for quantitative measurements**. A mix of quantitative and qualitative indicators, which complement each other, is often most useful.[95]

Standard **indicators defined by other partners** such as the World Bank, or United Nations agencies, can be considered if relevant.

*Source: DG NEAR*

### *9.4.3. Step 3 – Identify sources and calculation methods*


*Description*

[94] European Commission, 2015b.

[95] Last tip adapted *from Strengthening project internal monitoring How to enhance the role of EC task managers*, DG EuropeAid, European Commission, June 2007.

Each indicator, qualitative as well as quantitative, must correspond to an **existing source**, be it a statistical source or an administrative one. There are five types of sources: [96]

- international statistics;
- national statistics: nation-wide and sub-national level
- national registers/administrative sources at national and local level;
- data collected as an integral part of intervention implementation;
- ad hoc surveys.

**International and national statistics** are usually better suited for long-term impact and context indicators than for outcome indicators. Though, they can reflect the impact of relevant policy measures taken, also with the support of international partners.

**National statistics** are suitable to act as sources for outcome and impact indicators.

**Administrative data and data collected as an integral part of an intervention's implementation** are usually well suited for output and outcome indicators. There are, however, profound differences between the two sources of data: in general, administrative data, e.g., those from **national registers** (for example the national register of enterprises or the data collected by competition authorities) can provide useful material to calculate outcome indicators. Usually, instead, the data collected as an integral part of an intervention's implementation is more adequate for process and output indicators.

It is necessary to thoroughly investigate **which organisation at national level collects the data and how it does so**. This holds for statistical data and for administrative data. It is all the more important for administrative data, which, after all, is not primarily collected for scientific or research purposes. Therefore, the way it is collected may not be fully spelled out. In addition, this data may suffer from often "invisible" differences in internal procedures across public sector organisations, even within the same country, non-compliance with instructions, or lack of validation of data.

Information on data collection helps **ensure credibility**, verify the possibility of **comparing indicators** across interventions and countries, explore which data can be **aggregated for reporting purposes** on financial assistance interventions and make sure that it is possible to **build time series**. Intervention revisions or the dates for verification of milestones and targets (2017 and 2020 for the present planning/programming cycle) are ideal moments to ensure that data collection organisation and methodologies have remained the same over time.

Special care must be used when including indicators based on **ad hoc inquiries**.[97] Launching ad hoc inquiries takes up valuable time and human resources away from implementing and monitoring efforts. While recognising that in some cases ad hoc enquiries are needed[98], at early stages of developing and using performance systems it is, rather, **advisable to expand the available sources of data by improving information systems within the public service and by improving statistical information systems**.

[96] Adapted from Barca, McCann, 2011: 11 and European Commission, programming guide.

[97] There are reasons to advocate increasing the number of indicators based on ad hoc inquiries, especially in the name of expanding knowledge (Barca and McCann, 2011). Still, unless conditions are exceptionally favourable, it is best to concentrate efforts on building the indicator system and making it work with existing data, before introducing ad hoc surveys which require institutional capacity to routinely create and utilize knowledge which might not be yet in place.

[98] For example surveys to assess the level of political buy-in in a given area, the level of perception, etc.

Strengthening statistical compliance in partner countries[99] may support the effort of improving systemic collection of relevant data that can be used in the indicator system. This type of support includes activities aiming at:

- Enlarging the scope of statistical research within the statistical office of the partner country in a permanent way—for example, on gender-relevant data or poverty issues and
- Permanently improving the collection, processing, storing, and dissemination of administrative data in partner countries.

The **calculation methods** (which organisation processes the data and precisely which calculations are performed) and the method through which qualitative indicators are defined should also be made explicit. This is particularly relevant for indicators based on administrative data (i.e., data on individuals or enterprises benefitting from an activity) and for composite indicators.

**Composite indicators** should be used with care: they tend to obscure the phenomena they refer to and the underlying changes. They also require great care in their calculation. The best composite indicators are those that are already produced by recognised organisations (e.g., the World Bank's Doing Business indicators). Whenever it is necessary to use them, it is best to carefully define all their elements, to report carefully the sources of all the raw data, and to show very clearly the method with which they are arrived at.

**Timing considerations** affect the suitability of some indicators, which may be quantified with such delays that they lose the ability to signal change or whose **periodicity** does not fit in with the timing foreseen for milestones or targets. Timing has multiple dimensions.

- First, **collection, processing, and validation of data take time**: there is a delay between the moment in which the data necessary to build the indicator is available and the time it refers to. Time is necessary for the organisation collecting and processing the data to release it (in the case of data which is publicly available) or to make it available (for example in the case of intervention-generated data or administrative data). Often, delays may amount to one or two years. This may mean that information about a change that happened in, say 2016, may be available only in 2018. In addition, it is unlikely that in a sector all the indicators have the same delay.
- The second issue refers to **frequency**: **how often data is available**. Intervention-generated data such as expenditures may be accessible for partner countries and for DG NEAR offices almost continually. On the contrary, statistical data or data coming from systematic surveys may be available much less often, sometimes not even every year.

**BOX 24: EXAMPLES OF SUPPORT QUESTIONS AND TIPS IN IDENTIFYING THE SOURCES OF VERIFICATION**

The following set of **questions**[100] provides orientation:

- *Which are the available sources of information?*
- *What is their quality?*
- *How well do they fit with our needs?*
- *Which is their relative cost (in terms of time, money) of acquisitions and reliability?*
- *How often the data are made available?*
- *What are other donors doing?*
- *Is there an appropriate balance between quantitative and qualitative information?*

[99] "Under IPA II, countries are striving to improve statistical compliance. Countries would do well to direct the corresponding resources not only to Statistical agencies but also to line ministries in order to improve administrative data and to develop independent means of verifying indicators." (World Bank, 2014:17).

[100] The last five questions come from *Strengthening project internal monitoring Op. Cit.*

- *Are responsibilities for information collection clearly identified and understood?*
- *Are the existing formats for information recording and reporting adequate and are users clear about how to use them?*
- *Where are the most significant information gaps?*

**TIPS**

**Use data related to existing and credible sources and methodologies when building indicators.**

It is worth using data from reliable national administrative/statistical existing sources. This would be an added value for development of national monitoring and reporting systems.

Other sources and methods might be costly and time-consuming and in some cases this approach is not recommendable. In addition, this may threaten the accuracy of the results.

**Put yourself in the shoes of the person collecting the data when indicating the references to sources**. Make precise references to sources, in order to minimise the effort to locate the information—this is also useful for users of the indicators. Indicate the name of the database and all elements necessary to retrieve the data. Point out the exact point in reports where the information can be found. Avoid generic references such as "Annual Activity Reports."

*Source: DG NEAR*

### *9.4.4. Step 4 – Identify baseline values and timing*


*Description*

As mentioned elsewhere, an indicator provides information on the direction **and extent of a given change**. In order to do this, its value needs to be compared with other values, among which the **baseline**— the reference value at the start of the intervention against which progress will be assessed.[101]

The **description of the indicator** is key to understand which value to attribute to the baseline. For **output indicators**, which are defined in such a way that they refer to a specific activity (e.g., number of staff trained on M&E issues in the framework of the current intervention, kilometres of roads built with support from ENI/IPA II in a country), the baseline will normally be zero. On the contrary, if the output indicator is defined as referring to the situation in a field (N. of staff having received a training on M&E issues, regardless of the source of funding of training; kilometres of roads renovated each year by the road authority or a county government authority), it will be possible to identify a baseline.

Identifying a baseline is often easier for **outcome indicators** (e.g., the time-lapse between the beginning and the settlement of litigation in a civil tribunal).

The difference of the value of an indicator from the baseline indicates different things depending on the type of indicator:

[101] It is important not to confuse the value the indicator assumes before activities start with the value that the indicator would assume without public support through the activity (which can only be estimated through a carefully construed counterfactual). Especially for some activities (such as training or support for enterprises), there might already be analogous activities funded by other sources or by the precursor of the intervention of interest.

- in the case of **process and output indicators**—with data coming from administrative systems specific[102] to the activities of interest—the change of the value can be attributed to the intervention since outputs are within the sphere of control of the implementing partners.
- In the case of **outcome** indicators, instead, only evaluations can assess whether and to which extent the change is to be attributed, among other, to the intervention.

In the case of **qualitative indicators, baseline values should take the form of text,** such as "legislation not yet drafted". They are not to be quantified. For example, if for a qualitative indicator a '0' baseline is set, this would mean that its final value would need to be '1', which means nothing. **Quantifying qualitative indicators would result in losing the overview of the actual scope and quality of performance**. This being said, in some cases qualitative indicators can be quantified, mainly in aggregation exercises to facilitate comparison.

**The baseline value must refer to the closest possible moment to the beginning of the intervention**. Most indicators, however, need substantial time (one-two years) in order for the data to be collected, cleaned, processed, validated, and made available. That's also why it is preferable, as mentioned under Step 3 here above, to use indicators with existing and reliable sources of information.

**The time to which the baseline value refers to must be explicit**.

The baseline is the mandatory and the most essential element of the indicator tracking. It helps to overview the progress in attaining the results in the intervention.

**BOX 25: EXAMPLES OF SUPPORT QUESTIONS IN IDENTIFYING BASELINES**

The following set of **questions** provides orientation:

- *Given the definition of the indicator, does a baseline value equal to zero make sense?*
- *Which is the most recent data available?*

**TIPS**[103]

**Are the information requirements clear?** Before embarking on collecting baseline information, the key indicators should have been established and agreed. Otherwise the baseline exercise may lose focus and end up collecting information of little or no relevance.

**Minimum information.** What is the minimum information that would be useful? It is often best to be very selective, and focus only on the minimum information requirements that will be useful. More information adds complexity and cost. The capacity of the implementing partner to engage in baseline information collection, and to use the resulting information, needs to be carefully considered.

**Who is interested?** This is a key question, because if it is only the donor who has an interest, then it is unlikely that the information collected will be productively used by implementing partners. The baseline then becomes a largely academic exercise, with little prospect of local ownership or capacity building value.

**Look to existing sources first.** There is often a considerable amount of information already available through existing sources – it may just not be assembled and summarised in a user friendly and accessible form. Administrative records kept by government agencies at local levels, data from statistical offices, reports from research/academic institutions and other donor agency studies will often already contain useful and relevant data. Don't assume that primary data collection is required, and make sure that other interested donors are consulted and appropriately involved.

**Do you need to find out what people think?** If there are key indicators related to people's opinions or

[102] Usually, in partner countries, data pertaining to UE-funded operations are collected and stored separately from those referring to other operations. Sometimes, however, this may not be the case.

[103] *Strengthening project internal monitoring Op. Cit.*

attitudes, then it is likely that opinion surveys of some kind (or other communication processes) will be required. Opinion surveys can be extremely useful sources of information, however the quality of survey design, administration and the subsequent analysis/interpretation is critical to quality.

**What is the likely cost and benefit of different baseline options?** There will always be more than one option for collecting the required baseline information. Each will have different costs and benefits. The most appropriate approach will need to balance the objectives, with methods and cost. The theoretical ideal is almost never feasible or appropriate.

**Incremental approach.** It may be best to take an incremental approach to collecting baseline information.

While the theoretical ideal may be to have all relevant information collected at one point in time, this may just not be possible. The incremental approach may also have the benefit of allowing local partners to be engaged as part of a capacity development strategy. So while there may be no 'one point in time snapshot' – before the end of the intervention the capacity of the implementing partner to collect and use baseline information may itself have been developed, and be sustained into the future.

*Source: DG NEAR*

### *9.4.5. Step 5 – Establish targets*


*Description*

**Establishing targets** for indicators is the most delicate and difficult task in building up an indicator system. Based on technical knowledge about the intervention and its technological aspects—and about past experiences with similar interventions—this task incarnates decisions on the future (e.g., on the amount of financial and human resources which will support the intervention, and on the attention it will be able to count upon) and incorporates educated guesses on the environment in which it will unfold.

A target signals **how much change is expected and in what direction**. A target set for a time before the end of the intervention is a **milestone**—it indicates **how much road has been covered** and how much is still left to cover. In principle, different indicators might be used for milestones and for final targets, reflecting the sequencing of activities within the intervention. It is, however, **advisable, whenever possible, to use the same indicator for expressing milestones and final targets, since it simplifies data collection and interpretation**.

As in the case of baselines, **target values of qualitative indicators should take the form of text.** This being said, in some cases qualitative indicators can be quantified, mainly in aggregation exercises to facilitate comparison.

**Targets provide information to multiple stakeholders**—information that might be used in the future to **hold European Commission Delegation staff and intervention staff in partner countries accountable** and to **make decisions on the intervention**. The incentives are, therefore, mixed and even perverse. Overly ambitious targets might support negotiations initially, but might hinder the unfolding of the intervention later, should problems arise. Setting targets at too low a level is not defensible in the beginning of the intervention, and, in addition, provides little incentive to implementers to strive for results. Striking a right balance is of course especially sensitive with regard to indicators triggering the disbursement of tranches in the framework of budget support operations.

In principle, targets should not be set on an annual basis, but, rather, mainly in the case of outcome and impact indicators, **for key points in the planning/programming cycle**—for the 2014-2020 for the present planning/programming period, for 2017 and 2020 (this is currently the case for IPA II). In the case of long-term impact indicators, target **usually exceeds the end of the planning/programming period**.

At output, and depending on the intervention's duration at outcome level, the final target value will be set **when the intervention is completed**.

There are various elements, which help defining targets:

- The target must be attainable with the resources available and within the allotted time. A specific issue relates to **target values that are embedded in regulations, directives, or laws**, such as the objectives for ambient air quality assessment[104]. Although these are binding, they may still lay beyond the possibility of the country to reach them within the allotted time or with the available resources. In this case**, it is better to define the target at a lower level than the one mandated by law**. If the target is ambitious relative to the current situation of the country, but it is realistic, it can provide a better incentive to improve than a target that is impossible to reach. This is all the more important for indicators that may be taken in consideration for performance reserve instruments.
- **Past trends**: past experience shows what is likely to be attainable and what is not. It is not possible to just extrapolate past trends into the future: it is necessary to take into consideration the amount of resources devoted to the activity (and compare it with the resources available in the past from various sources), to examine how the activity has changed, and take into consideration also changes in intervening factors. For example, "indicators such as net employment growth are largely driven by GDP growth and activities in other sectors (real sector). To the extent that good projections exist for these variables, they can be used to develop different scenarios that translate into a target range. Certain indicators also have an implicit range for potential improvement- for example, what is the highest potential year-on-year productivity growth in a sector?[105] For soft areas of intervention target setting is particularly difficult, especially while dealing with sensitive reform areas (like Rule of law), where no mechanical model can help and results might depend on the difficulty of implementing a reform in a particular conjuncture of the political cycle/environment.
- **Other sources**. It is necessary to take into consideration other policy efforts (by national authorities or international funders) in the same sector and partner country: which is the potential of these policies? Are they going to multiply or depress the effects of the EU support?

**BOX 26: EXAMPLES OF SUPPORT QUESTIONS AND TIPS IN IDENTIFYING TARGETS**

The following set of **questions** provides orientation:

- *Based on past experience or on available knowledge, which is the level of output reachable with the allotted resources?*
- *If the activity has been implemented before, what has changed in the implementation system?*
- *Do the changes in the implementation system allow for higher or lower levels of outputs?*
- *Which are realistic assumptions on the outcome level that is possible to reach given the expected level of output produced?*
- *Which evidence supports the hypothesised link between outputs and outcomes?*
- *Is it possible to only detect the direction of change?*

**TIPS**[106]

**Involve the right people.** As with the selection and use of indicators, it is important that those responsible for implementation on the ground take the lead in the setting of targets. Ownership of targets by implementing partners is critical. Donor driven target setting does not have a good track record. Similarly, targets (as well as indicators) should be aligned to the extent possible with partner priorities and systems.

[104] Annex I of Directive 2008/50/CE.

[105] The World Bank, 2014: 31.

[106] *Strengthening project internal monitoring Op. Cit.*.

**Targets need to be realistic.** If targets are not realistic (given the operating context, intervention scope and resources available), they will quickly become irrelevant, are likely to be ignored and/or may simply cause frustration. Choosing appropriate targets requires adequate background research and consultation. They should not be 'plucked from the air'.

**Targets need to be reviewed.** Targets need to be regularly reviewed, as part of the ongoing monitoring and review process. This helps ensure their continued relevance.

**Need incentives to report accurately and honestly.** If targets are not met, then the messenger must not be shot! The reasons for the target not being met must first be assessed in an open manner, without initially apportioning blame. The reasons that a target is not met may have nothing to do with the competence of intervention's management.

*Source: DG NEAR*

### *9.4.6. Step 6 – Validate indicators*


*Description*

Ideally, as mentioned above under Step 2, identifying indicators, and especially establishing indicators, should be done in a **participatory way**, including relevant stakeholders, especially the individuals in partner country organisations who will be responsible to implement activities and those having to collect data and transfer them. In reality, however, this may not happen. This contrast between an "ideal" situation and the reality of processes creates an opportunity.

**The validation step allows programmers to use indicators in order to re-discuss the entire intervention** with sector experts, statistical offices, implementing agencies, and partner countries officials, experts, and socio-economic partners, including representatives of potential beneficiaries. **This may result in improvements and in a deeper sharing of intervention goals, objectives, and especially mechanisms and policy instruments than would have otherwise not happened**. At this step, various instruments may be used to bring together relevant stakeholders: conferences, seminars, or training, mixing individuals from various organisations from both the European Commission (Headquarters and Delegation staff) and the partner country.

**The "technical" goal of validation is to verify whether sources are actually available, whether the indicators may provide timely information, whether they are politically acceptable**. It also constitutes an opportunity to assess the **appropriateness of descriptions** and, in general, **to assess to which extent each indicator responds to quality principles** (refer to § **9.2.3 What are the characteristics of indicators?**). The same principles used to generate relevant questions in Step 2 might be used in a different sense for validation: while the same questions (e.g., on the way the data is collected and processed) need to be asked again, the collective nature of the effort brings about further issues.

**Box 27: Avoiding resistances in indicator's validation**

**Winning resistances** is particularly important. Delays in and neglect towards building an operating performance indicator systems result from resistances against a practice that primarily aims at increasing accountability. There are many tools, which can be used to lower these resistances:

- It is necessary to create an **actual demand for data**.
- It is important not to overly burden the organisations and individuals, which have to provide the data.
- An **internal drive** is needed: organisations should be allowed and even encouraged to access data from

the monitoring system and use it. Data collection is therefore decentralised (beneficiaries/partners provide data, insert data in the monitoring system, collect information from existing sources, and process the data to build indicators) and the utilisation of data is centralised.

*Source: DG NEAR*

Please refer to Annex 7 for identification of validation questions.

Validation may also provide opportunities for further tasks.

- First, **stakeholders have the opportunity to earn a deeper understanding of what to expect from the intervention**. In turn, this might result in **better implementation and in greater surveillance over the entire system**.
- Second, implementers (or at least some of them)[107] **learn about the metrics which will be used to measure their performance and which they contribute to create by transferring data**. If the validation is conducted in a participatory way, they are called to agree on indicators. **This may, in principle, result in better and more timely data collection**.[108]

### *9.4.7. Step 7 – Design the data collection systems*


*Description*

In order to ensure the quality and timeliness of monitoring data, it is necessary to **carefully set up a data collection and processing system**. The system must be:

- **Sustainable:** sufficient resources must be continuously available over a long period and
- **Resilient** to shifting priorities both in the interventions and in the partner country.

These points cannot be stressed enough. Building a data collection system includes taking care of issues such as **good record keeping** within both the public sector and the private organisations involved in implementation, the **reliability and timeliness of data delivery**, and the **credibility** of the data. This involves **foreseeing conflicts** that may arise and **finding out ways to dispel them**: implementing partners are the ones responsible for providing data about outputs and sometimes outcomes[109]. They may feel that data are extracted from them, and resist. **Resistance rarely takes the form of open refusal**. More often, it takes the form of **administrative neglect**: data collection becomes a very low status activity and/or is felt as an imposition, taking away time from more valuable tasks. Implementing agencies may fragment the tasks (e.g., collecting data and inputting them) and attribute some of them to very junior or unqualified resources.

[107] This only applies to a limited number of implementers and beneficiaries of the activities. Many of these are identified only much later during implementation. This is one of the reasons why delaying these steps may actually be beneficial: the more complete the intervention is, the more stakeholders will have the opportunity to participate in key processes.

[108] It is important, however, not to overly rely on these mechanisms. The operation of an indicator system basically depends on the actual pressure there is to obtain knowledge on performance and on whether there is somebody using the knowledge or not. In turn, these elements depend on whether top management (at the European Commission as well as in partner countries) demand data, on whether the system is so open that the public can effectively demand data, and on whether the organizations which have the task to fill in the data have the right incentives to do so.

[109] As the Better Regulation mentions, evidence providers will vary according to the intervention and the indicator. Sometimes evidence is needed from individual stakeholder (e.g. businesses or citizens); sometimes it can come from a collective body (e.g. business or consumer organization, local/regional/national level). Every effort should be made not to impose excessive burdens on the stakeholder in terms of monitoring requirements.

**BOX 28: BASIC 'QUALITY' ELEMENTS ANY MONITORING SYSTEM SHOULD INCLUDE AND CLEARLY DEFINE (AND DOCUMENTED)**

- Objectives and Principles, to guide the approach to monitoring.
- A well-defined intervention logic, with clear result statements and associated indicators
- Baselines showing the situation at the beginning of the intervention
- Targets/quality standards against which performance can be assessed.
- Information sources and collection methods.
- Formats and procedures for data collection, recording, analysis and reporting.
- Monitoring tasks, roles and responsibilities.
- Consideration should be given to who will have responsibility for gathering data and who will be responsible for providing it. Evidence should be gathered at appropriate level with consideration to cumulative burden it could trigger.110
- Frequency of measurements
- Timing of the evidence gathering needs to be considered vis-à-vis progress of intervention's implementation and reporting requirements. [111]
- Reflection, review and decision making processes
- A communication plan; and
- Resources and budget for monitoring and reporting

*Source: DG NEAR, adapted from Strengthening intervention internal monitoring How to enhance the role of EC task managers, DG EuropeAid, European Commission, June 2007*

**Collected data needs to be verified for mistakes**, both casual (for example typing mistakes) and systematic (for example, when there are misunderstandings on exactly what is being measured, at which moment, and with which unit). This work requires dedicated and trained internal human resources, emphasis from management providing the drive to periodically perform the verification, time for individuals to verify the data, and a deep knowledge of activities, as well as the ability to cooperate with the individuals performing the data collection. **Verification is, therefore, a task that is best performed internally**.

**The most important source of data verification, however, is utilisation of information from indicators for demonstrating results, reporting, and research**. It is when trying to use the data that individuals are at the same time forced and willing to verify its accuracy and credibility. It is when processing data that anomalies and mistakes stand out. This is a powerful argument in favour of **making data from indicator systems as widely available as possible**.

**BOX 29: DG NEAR'S DATA COLLECTION SYSTEMS**

In the framework of DG NEAR, the progress in achieving indicators will be tracked through DG NEAR's Management Information System (**MIS**) that complements CRIS/Iperseus[112] financial information[113]. A new module in MIS will report on the progress of the IPA II indicators (still to be developed for ENI indicators). The system will allow aggregation of the indicators at country, sector and instrument level and thus will serve the purpose of a main monitoring tool for tracking indicators.

In the coming years, it is expected that a new IT tool, the Operational Information System (OpSys), will help improving the efficiency of the management and implementation of the EU external action interventions, which

[110] European Commission, 2015b.

[111] Ibidem.

[112] This system has been put in place for decentralised managed projects in the Enlargement countries.

[113] Information and guidance on how to encode indicators in MIS is provided in the ARES note n. 3842820, dated 22/07/2016.

includes the phasing out of CRIS. It will help EC staff to report on results and as such allow informing on indicators.

*Source: DG NEAR*

**And finally, it is of utmost importance to clearly define monitoring tasks, roles and responsibilities for data collection, analysis and reporting**. In this regard, it is important to ascertain whether the necessary technical skills required for monitoring are available both in the implementing partner and in EC HQs and EUDs[114].

**BOX 30: EXAMPLES OF SUPPORT QUESTIONS IN DESIGNING A DATA COLLECTION SYSTEM**

The following set of **questions** provides orientation:

- *How does the data collection system work in each of the implementing organisations and within the partner country public service?*
- *How independent is the statistical system and how collaborative it will be?*
- *Would a public source be considered more reliable than a private firm contracted out to collect the data?*
- *Is it necessary to ensure that multiple organisations sit on a committee reviewing the data? In this case, are sufficiently skilled professionals available within the organisations?*
- *Which organisations use the data? For what do they use it? Which feedback do they provide on data quality? Is there a procedure in place to collect this feedback? What would be required to make the data open?*
- *Who undertakes analysis of the available data and information and at what level within the reporting hierarchy?*

*Source: DG NEAR*

**BOX 31: MAINTAINING AN INDICATOR SYSTEM**

For all its complexity, identifying indicators is far from being the most difficult task: the real challenge for an organisation and its partners is, rather, to systematically operate the data collection over time, as priorities change and interventions evolve, and as everybody's attention is diverted to spending the financial resources and to overcoming the obstacles to implementation.

Many ingredients, from both the EC and partner countries' sides, must be present over a long period, often largely exceeding the planning/programming period:

First, a sustained and continuous external drive to collect and use data. Ideally, top management in both the European Commission and in the partner countries should provide this drive, by systematically demanding for demonstration of results of interventions. Although there seems to be a progress on this, requests for data on intervention results remain episodic.

Second, it is necessary that this demand is continuous and expected from both EC management and partner countries, in order to keep organisational attention focused on continuously producing reliable data.

Third, the necessary resources must remain in place for the entire time the monitoring lasts—which largely exceeds the planning/programming period. At a minimum, resources include skilled people whose time is at least partially devoted to data collection and interpretation; organisational resources to collect, store, and recover data from administrative sources in an easy and reliable way; procedures which allow for sufficient advance time to integrate data from indicators system—especially output and outcome indicators—in the narratives included in periodic reports; and sufficient support provided to statistical offices and the activities

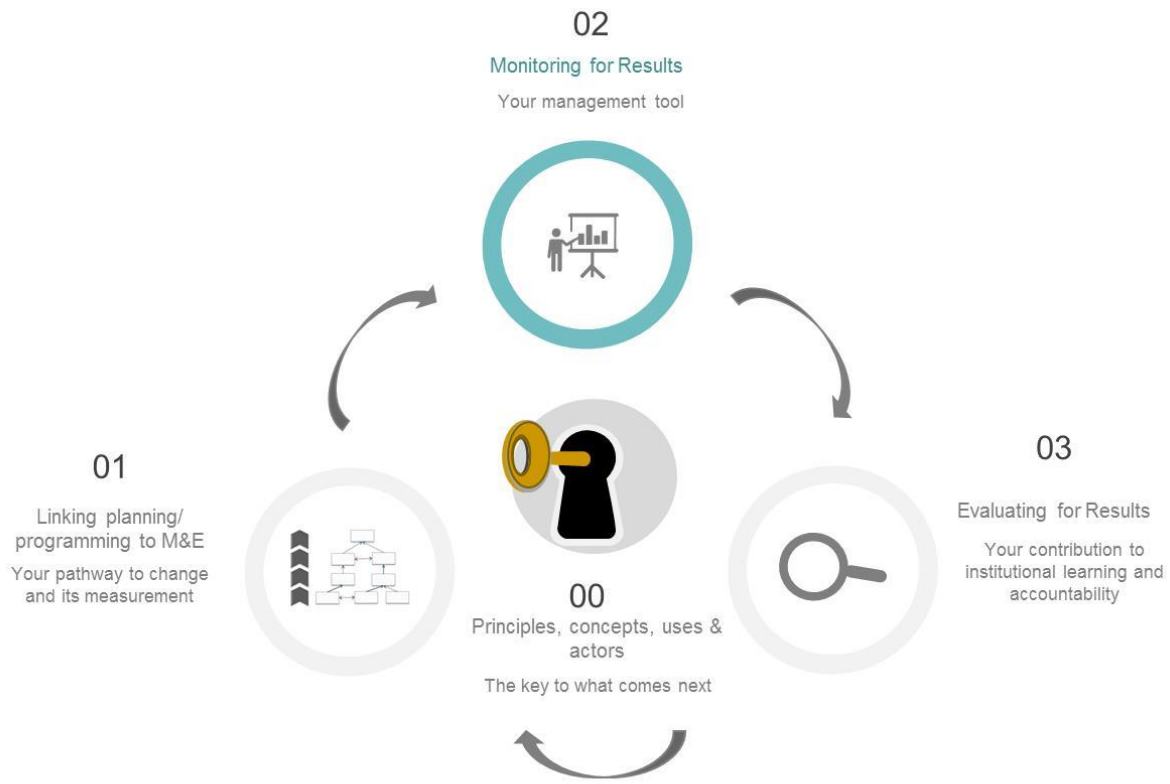
[114] An institutional capacity assessment checklist, to be adapted to each specific case, can be found in *Strengthening project internal monitoring Op. Cit.*

Also, the Rapid Assessment for Capacity Development (RAC) methodology (http://capacity4dev.ec.europa.eu/public-cd-tc/minisite/rapid-assessment-tool-capacity-development-rac ) can be used (with some adaptations).

they perform.

Fourth, all stakeholders should maintain trust in the data from the indicator system. It is perfectly acceptable, and even desirable, that indicators data raise more questions than they answer, thus requiring ex-post evaluations and further studies. Nothing, however, is worse than having stakeholders challenging the credibility of data once they are available and used. It is necessary, therefore, to set up ways to minimise the risk of such a challenge.

*Source: DG NEAR*




# PART 2: MONITORING FOR RESULTS

**BOX 32: SUMMARY OF KEY MESSAGES ADDRESSED IN PART 2**

***Monitoring starts from a proper risk assessment. Implementation of (both internal and external) monitoring activities depends on it.***

***The monitoring plan needs to clearly state the type of monitoring activities. This is important so that complementarity and synergies can be found with other monitoring activities and related actors.***

***The Monitoring exercise must be focused on the real corporate information needs. Indicators defined during planning/programming are expected to be reported on during internal and external monitoring. Depending on the needs (feed decision-making process at the management level, learn and/or accountability on the use of resources), a choice needs sometimes to be done on what to monitor. This is particularly the case of internal monitoring done by EC/EUD staff.***

***Lessons learned play a key role in data analysis and interpretation (and subsequent reporting): they are the basis of corrective measures.***

***They are to be considered as outcomes of a learning process, which involves reflecting upon the experience.***

*Source: DG NEAR*

## 10. TYPES

It is possible to distinguish among three broad categories of monitoring:

1) **(Action's) operational monitoring**[115], which can be internal or external:

   a. ***Internal monitoring, which is the main object of the present Part,*** is implemented both by implementing partners[116] [such as agency's staff, government's personnel[117], other donors, non-state actors (private sector companies, NGOs, etc.)] and by EC HQ and EUD staff.

[115] The most recent ROM results for ENI showed that well-structured M&E frameworks and approaches positively impacted on efficiency and effectiveness. In some cases, even simple frameworks and informal meetings were sufficient.

[116] Implementing partners refer to the organisations that have direct responsibility and authority for project implementation, including management of the available resources, implementation of activities and achievement or results. An intervention may have one or more implementing partners, and such partners may include government agencies, non-government organisations and/or private contractors, *Strengthening project internal monitoring Op. Cit.*

[117] National authorities are normally invited to monitor action implementation. This is the more important, the higher their involvement in the implementation of financial cooperation. Whilst these responsibilities are marginal under complete centralised management, where Commission services (EUD or HQ) are responsible for implementation, they become more relevant when national authorities have management responsibilities, and especially under complete decentralised implementation system. More specific guidance is provided in Annex 16 for IPA II, and it will be complemented by the provisions referred to in the so-called entrustment process.

i. Implementing partners' monitoring collects and analyses data to inform on progress towards planned results' achievement to feed decision-making processes at the action's management level and to report on the use of resources.

ii. EC HQ and EUD staff monitoring aims at complementing implementing partners' monitoring, especially in key moments of the actions cycle. It also aims at ensuring a sound follow-up on external monitoring recommendations and at informing EC HQS/EUD management. EC HQ and EUD staff is to be considered as 'informed dialogue partners'.[118]

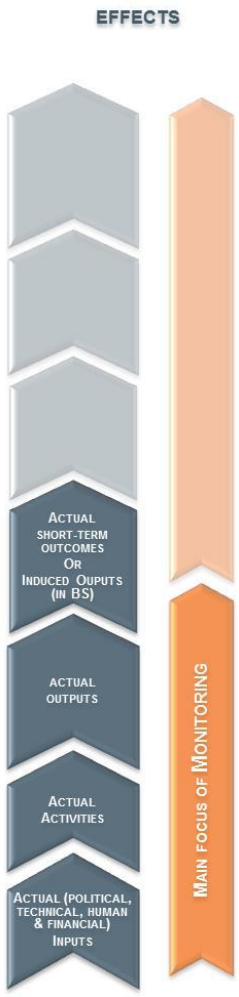
Both types of internal monitoring are meant to inform and provide support to external monitoring.

b. ***External monitoring***: external monitoring in the practice of the external relations at the Commission is implemented by means of Results Oriented Monitoring (**ROM**)[119]. This has been introduced by the EC to get an independent view on action's performance. External monitors assess performance based on some of the OECD DAC evaluation criteria. The new approach to ROM aims at providing support to EUDs/HQ services in action monitoring and reporting functions and is regulated by a specific Handbook regularly updated by DEVCO.[120] & [121] Its focus is presently on high-risk actions.

Beside the ROM review on on-going actions, an end of action results reporting has been introduced to support EC HQs and Delegations[122] in identifying and checking the most relevant results on their respective actions to report on.

Another form of external monitoring is the one foreseen in the framework of budget support operations, when experts are recruited ad hoc in order to monitor the implementation of the reform agenda and typically help the Delegations in assessing the state of implementation of indicators prior to disbursement decisions.

**FIGURE 13 : MONITORING, ITS PLACE WITHIN THE INTERVENTION LOGIC**




*Source: DG NEAR*

[118] *Strengthening project internal monitoring Op. Cit.*

[119] Two contracts are managed by DG NEAR: one for Neighbourhood countries; the other for the Western Balkans, Turkey being subject to ROM under a dedicated contract in decentralized management. The ROM Module, in which the ROM reports are made available, is available at: https://myintracomm.ec.europa.eu/dg/devco/resources-procedures/it-information-systems/rom/Pages/index.aspx.

[120] https://myintracomm.ec.europa.eu/dg/devco/quality-impact/monitoring-rom/Pages/external-monitoring.aspx.

[121] The main changes introduced in 2015 concern the use of specialised thematic and sector expertise (i.e., senior and medium sector experts are prioritised over M&E experts); longer field work (7 days instead of 5) for stakeholder consultation; and stronger quality control arrangements (dedicated Quality control expert per action). Impact is no longer addressed. Also, the past rating system has been replaced by a system of traffic lights. The specific ROM manual can be found at: https://encrypted.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwjQw8HE7PfMAhVCUBQKHbxMBdgQFggbMAA&url=https%3A%2F%2Fcapacity4dev.ec.europa.eu%2Fsystem%2Ffiles%2Ffile%2F28%2F10%2F2013_-_1159%2Fec_rom_handbook2012_en.pdf&usg=AFQjCNF9tTxzoQujnrYG-TdJJpHk8JvDXA.

[122] Ex-post monitoring does not exist any longer according to the new approach, though it has been maintained for Western Balkan countries, where past experience had been positive and there was an increasing demand.

### 2) Overall sector strategic monitoring

Strategic monitoring can be organised, whenever relevant, to ensure a proper, joint supervision of complex, interrelated interventions implemented in a given sector. Ideally, this monitoring will take place at a higher, more strategic level. This is the case of:

- Sector Monitoring Committees under IPA II
- Inter-service meetings at sector level encompassing different dimensions of interventions (sector strategies, activities implemented by the Government itself together with different development partners).
- Specific donor coordination groups set out around a budget support programme. Their role will be to discuss in an holistic way issues related to the implementation of sector reforms; the implementation of interventions' activities; the review of specific requirements foreseen by the programme (like fulfilment of conditionalities and/or degree of realisation of targets in budget support operations); review of relevant sector indicators.
- Sub-committee meetings in the framework of the relevant partnership agreements to follow-up at sector level on reform agenda and bilateral relations.

### 3) System monitoring

By system monitoring we refer to the review (made by means of dedicated meetings, supervision missions, performance audits, reports, etc.) that the different actors (Commission, national authorities) put in place to assess the way in which the whole system organised in relation to the implementation of financial implementation takes place. This is especially relevant and structured in the framework of IPA II, where a specific IPA Monitoring Committee is foreseen[123].

**TABLE 12: MONITORING ACTIVITIES AND RELATED ACTORS**

| | Strategic | Operational | Contractual/ Financial |
|---|---|---|---|
| ***1a - Action's monitoring – internal*** | | | |
| *Meetings with contractors (implementing partners)* | - | EUD (in centralised management), Line Ministries (in decentralised management | Contracting Authority (CA) (EUD in centralised management) |
| *Action Steering committee* | National coordinating authorities, EUD, line Ministries, contracting authority, contractors, beneficiaries, stakeholders | | |
| *Internal monitoring missions* | - | EUD, National coordinating authorities; line Ministries | CA |
| *On the spot checks* | - | EUD | EUD/CAs |
| ***1b - Action's monitoring – external*** | | | |
| *ROM* | - | Contractor | - |
| ***2 - Strategic monitoring*** | | | |
| *Strategic sector meetings* | National coordinating authorities, CAs, line Ministries, EC, stakeholders | | |
| ***3 - System monitoring*** | | | |
| *Supervision meetings* | - | - | National coordinating authorities; EC HQ |
| *Performance audits* | - | - | National coordinating authorities; EC HQ |
| *Supervision missions* | National coordinating authorities, CAs, line Ministries, EC, stakeholders | | |

*Source: DG NEAR*

[123] See Annex 16.

These monitoring activities are informed by relevant data production and gathering to be ensured by the relevant entities: i) International institutions for impact/context indicators, ii) National administrative services and statistical offices for outcome indicators, and iii) monitoring information systems at Intervention level for output indicators.

## 11. MONITORING MANAGEMENT PHASES: PLANNING, IMPLEMENTATION, REPORTING, FOLLOW-UP

### 11.1. What to monitor and resources allocated

#### *11.1.1. Risk assessment*

Though a certain degree of operational monitoring (like screening of reports, attendance to Steering Committee meetings…) needs to take place anyway, an appropriate approach in planning monitoring activities, also for cost-efficiency considerations, is to start from a proper risk assessment.

Before the risk assessment takes place, a list of all on-going actions needs to be set out.

Implementation of (both internal and external) monitoring activities will be normally based on the risk assessment made at Action and contract level.[124] This will normally reflect dimensions related to:

- The political, physical, economic, etc. context in which the activities take place
- The institutional environment: the actors involved in the activities and the interrelations existing among themselves
- The technical dimensions of implementation of the activities
- The procurement and related legality/regularity issues
- The financial volume of the action/contract.

**The risk assessment represents a key ingredient in the monitoring cycle**. Though several approaches can be used to do the risk assessment, DG NEAR staff is requested to make the risk assessment on the basis of the approach that has been developed and integrated in the Management Information System (MIS). This system refers to four basic areas:

**TABLE 13: THE FOUR AREAS OF A RISK ASSESSMENT**

| Risk type | Risk description | Likelihood (from 1: low, to 5, very high) | Impact (from 1: low, to 5, very high) | Risk (from 1: low, to 5, very high) | Justification |
|---|---|---|---|---|---|
| *External environment* | *The political or institutional background plays an important role for the achievement of the objectives of the action and there is a risk arising from beyond the DG/Commission. The action needs to be monitored carefully on those aspects* | | | | |
| *Planning, processes and systems* | *There is a risk that the impact foreseen is not achieved due to such issues, for example because the structure of the action is complex. The action will require particular monitoring of aspects related to planning, processes and systems.* | | | | |

[124] This will normally reflect dimensions related to the environment (political, physical, economic, etc.) in which the activities take place, to the institutional environment (the actors involved in the activities and the interrelations existing among themselves); the technical dimensions of implementation of the activities; the procurement and related legality/regularity issues; the financial volume of the project/contract. In DG NEAR practices, risk assessment is made according to a specific module developed in the MIS.

| Risk type | Risk description | Likelihood (from 1: low, to 5, very high) | Impact (from 1: low, to 5, very high) | Risk (from 1: low, to 5, very high) | Justification |
|---|---|---|---|---|---|
| *Legality and regularity* | *There is a risk that the rules and obligations under the contract may not be respected. There is exposure to fraud and corruption, or issues with the financial capacity of contractor - need for robust audit / control systems.* | | | | |
| *Financial exposure* | *The higher one amount of money is committed in one single action, the higher is the financial exposure for the EC to lose those funds* | < 1M€<br>>= 1M€<br>>= 5M€<br>>= 10M€<br>> 20M€ | | | |
| *Global risk* | | | | | |

*Source: DG NEAR*

The first two areas, referring to operational aspects, beside the risk exposure, represent key elements that, if assessed as very risky, should lead to a sustained planning of monitoring activities, whilst mostly, but not only, legality/regularity issues can lead to the proposal of an audit.

Ideally, the object of a monitoring exercise should be the set of activities put in place in order to attain the predefined objective. It can best focus on:

- The whole action: a project, a programme or a Budget support operation (including the tranche transfer and the complementary activities);
- A specific element of the action, when it is articulated in different components/dimensions each focusing on a different specific objective;
- Specific contract(s): in case of grant schemes, made up of an articulation of different specific projects; twinning programmes (articulated in specific twinning projects); articulated infrastructural actions, with independent works located in different areas). This can also be the case of contracts requiring specific control activities.

The risk assessment should first take place at Action level, and then at contract level, which is the level also required to come up with the audit plan. According to the above considerations, the optimal focus will be identified (Action, specific contracts or group of contracts).

Other considerations (like innovation, political sensitivity, lack of sector expertise, lack of budget/time for internal monitoring) can complement the risk assessment in justifying a monitoring activity. The typology of monitoring activity foreseen as well as the justification for a ROM (problematic, lack of sector expertise, innovative nature, politically sensitive) should be clearly stated in the MIS module.

The risk assessment will have to take place typically in the last Quarter of the year, consistent with the instructions received.

### *11.1.2. Monitoring plans*

A monitoring plan[125] needs to be prepared at the end of each year to reflect monitoring activities foreseen the year after. This should be the case for all EU Delegations and Offices and operational units, as far as internal

[125] Monitoring plan will be integrated with the evaluation plan since 2017, making use of a dedicated template.

monitoring is concerned. With regard to the ROM plan, this will be coordinated by DG NEAR HQ's Monitoring & Evaluation service.

The focus of the monitoring plan should reflect the considerations made in the paragraph above.

**BOX 33: STEPS IN DRAFTING AND MANAGING A MONITORING PLAN [126]**

**FIRST STEP: IDENTIFY INSTITUTIONAL RESPONSIBILITY FOR DRAFTING AND MANAGING THE MONITORING PLAN**

The first step is to identify a clear institutional responsibility for drafting and managing the Monitoring Plan and for keeping all relevant actors informed: in DG NEAR HQs operational Units, in Delegations and within the partner coordinating structure there must be a Monitoring Focal Point.

**SECOND STEP: IDENTIFY WHAT TO MONITOR**

The **starting point is the risk assessment**.

In addition to this, the Monitoring Focal Point can use the following **sources**:

- Regulations and other legally binding sources or instructions, as appropriate, which envisage monitoring;
- EC HQs and Delegations need to schedule ROM for **actions above €750,000.** Delegations may ask for ROM of particularly innovative actions, pilot actions, and sensitive actions (actions in crucial sectors of interventions, like Rule of Law, or for which there have been implementation problems or conflicts or about which there has been controversy during action preparation) or themes on which there is little evidence.
- Top management requests for information may signal that there is a need for evidence on an action, group of actions or theme.
- For Delegations, consultation with the national aid coordinator and with the partner country authorities.
- Actions about which there are, among other new strategic decisions, complaints from stakeholders, EU infringement procedures (European Commission, 2015c: 260).

These sources provide an indication of information needs. The Monitoring Plan Manager should determine which ones can be satisfied by using information from previous monitoring or from existing studies and evaluations to satisfy these needs. These should not enter the Monitoring Plan.

The Monitoring Plan manager should **prioritise** the possible monitoring. Criteria are:

- **Appropriate timing** in consideration of the stage of implementation (ROM should not take place in the first and in the last six months of the duration of the action)
- **Interest** in the evidence from the DG top management or the Delegation management
- **Usefulness** in relation of particular concerns and/or needs
- **Complementarity** with other monitoring activities performed by the national authorities or ROM
- For EC HQs/EUD staff internal monitoring missions and/or on the spot checks, **feasibility within the timeframe and resources available.** If it is impossible to fit the monitoring in staff's work programme, consider using the ROM.
- **Interest** from national aid coordinator, partner country's authorities and feedback from **stakeholders** (European Commission, 2015c: 260).

Please also refer to **TABLE 2: GUIDANCE TO THE SELECTION OF THE APPROPRIATE CONTROL TOOL.**

**THIRD STEP: CHOICE ON MONITORING'S TYPE AND ON AVAILABLE RESOURCES**

For each monitoring, the Monitoring Plan manager needs to collect information in order to make choices:

- Whether the monitoring is going to be conducted internally (an option only possible when there are available human resources and budget) or externally (which doesn't have a cost for the Delegation, in case of the ROM, but needs to be agreed upon with the ROM coordinator ).
- The amount of needed financial and human resources. For internal monitoring missions or on the spot checks, the Monitoring Plan should at least indicate a range, based on experience.

**FORTH STEP: SCHEDULING THE MONITORING EXERCISES**

[126] The management plan might be integrated at a later stage with the evaluation plan.

For each monitoring, it is necessary to **identify the time when results are finally available** (mainly in the case of ROM, but not only). From this, the Monitoring Plan manager can calculate the time when the preparation should start. The start date is calculated from the moment in which the preparation starts.

ROM should not be foreseen in the first six months since the start of the action, and in the last six months, unless there is the perspective of an extension or continuation.

Delegations should also take into consideration additional **milestones** specific to the partner country (e.g., the national planning/programming cycle).

**FIFTH STEP: SHARING**

Once the Plan is approved, the manager of each Monitoring Plan ensures its **sharing** vis-à-vis the national relevant stakeholders.. The consolidated monitoring plan should be available on DG NEAR intranet.

**SIXTH STEP: MANAGING A MONITORING PLAN**

At this point, the difficult part begins. Managing a Monitoring Plan implies many tasks. First, the Monitoring focal point needs to make sure that programme managers activate each monitoring according to the schedule. Especially when many months have elapsed since final approval of the plan, or if there have been significant changes in the environment, this requires ensuring that the information need which gave rise to the monitoring is still present and that it still cannot be satisfied through existing data. Then, the manager needs to check whether the allotted resources are still available.

The Monitoring Plan needs to be updated twice a year on the basis of the most significant results, formal changes in management conditions and problems identified, and suggested actions to be taken.

*Source: DG NEAR*

**BOX 34: CONTENT OF A MONITORING PLAN**

A monitoring plan must clearly state:

- The **responsibilities** for the plan: which entity (unit/Delegation) is responsible for drafting and updating it each year; for deciding in-year revisions of the plan, if necessary; for activating the units and individuals responsible for each monitoring; for ensuring that the resources are available and used;
- The **title** of the relevant Action/contract or group of contracts;
- The Decision CRIS number;
- The CRIS number of the contract(s) subject to monitoring;
- DAC sector code;
- The amount of the total number of contracts concerned;
- Name of the **contact** point for the plan;
- Name of the Operational Manager in charge for the implementation of the specific monitoring activity;
- The risk level at Action and/or contract, as appropriate;
- The **time frame** of implementation;
- The **type of monitoring activities** that will be undertaken in the following year (assessment of documents; attendance to Steering Committees, monitoring mission, OSC, ROM);
- For EC HQs/EUD staff internal monitoring missions and/or on the spot checks, an indication of the **resources** available for monitoring for each year. ROM has its own budget. And internal monitoring done by implementing partners has also its own budget (within the intervention's budget).

*Source: DG NEAR*

#### *11.1.1. Funding of the monitoring*

Consistency should be ensured between planning of monitoring and budget. This is especially relevant for internal monitoring, where mission, seminar/workshops, specific studies, etc. budget has to be secured to reflect monitoring needs.

Shortage of mission budget can be a reason to opt for an increased use of ROM missions, as appropriate. The Head of Operations and the monitoring focal point would have to interact with the ROM coordinator in that respect.

### 11.2. Implementation

#### *11.2.1. Readiness assessment of the internal monitoring system for a given action*

As mentioned under Practical steps in defining, verifying, using indicators for reporting for Action programmes, in particular Step 7, it is expected that a system tracking inputs, activities performed (as per relevant Action document), outputs produced, and preliminary outcomes achieved will be put in place at action level. This is to be done during planning/programming.

Even though readiness assessment of such a system needs to take place all along an action's implementation, it is particularly important to do it before the actual monitoring related activities take place.

#### *11.2.2. Decide what to monitor*

Operating the data collection during implementation is a time-consuming and costly activity. It is therefore of utmost importance **to focus monitoring exercises on the real information needs.** Indeed, monitoring is not intended to substitute other reporting moments.

As mentioned under Part 1, **monitoring** (and evaluation) is **framed within an action's result chain**. **As such, indicators defined during programming are expected to be reported on in monitoring** and to the maximum extent possible in evaluation exercises. In the case of monitoring, indicators to report on go from inputs until outcomes. And while it is expected that the internal monitoring system gathers and analyses data pertaining to each of the indicators defined during planning/programming, **it might also be possible that in particular moments, depending on the needs (feed decision-making process at the management level, learn and/or accountability on the use of resources), a choice needs to be done on what to monitor. This is especially the case for internal EC/EUD staff monitoring.**

**BOX 35: EXAMPLES OF SUPPORT QUESTIONS IN DECIDING WHAT TO MONITOR[127]**

The following represents a tentative list of questions one should address:

- Which activities consume the highest levels of (human, financial) resources?
- Is there a specific action's dimension that needs updated data and in-depth analysis?
- What are other key stakeholders' monitoring?
- What do you want your main stakeholders to act upon based on monitoring results?
- How will the main stakeholders modify their behaviour once they know that a given indicator is monitored?
- Which areas were monitored during the previous planning/programming period? In this framework:
    - which indicators were selected (in addition to those already included in your monitoring system)?

[127] The last three questions come from Strengthening project internal monitoring How to enhance the role of EC task managers, DG EuropeAid, European Commission, June 2007.

- How EC HQs/EUD/national authorities /other stakeholders did use the information?
- Which have been strong points and problems? How were the problems solved?
- What are the existing physical and financial resources available for monitoring?
- What is the level of staff skills and their understanding of what is required?
- Are these adequate?

*Source: DG NEAR*

### *11.2.3. Actual (internal) monitoring*

Intensity in internal monitoring depends on the management mode, activities at stake, costs effectiveness considerations, budget availability, risk and complementarity with ROM[128]. Also, it is important to highlight that internal monitoring should be complemented by other sources, such as ROM, or specific studies commissioned to analyse specific areas of concern/interest.

Internal monitoring consists of activities such as:

- Analysis and feedback on actions' reporting documents and data.
- Attendance to actions' Steering Committee meetings and reviews of budget support operations[129]. Tentatively, attendance to action Steering Committees should systematically take place for all management modes.
- Meetings. Information sharing and discussion, focused on problem solving and forward-looking planning, need to be promoted and facilitated. This usually takes place during formal or informal meetings. They normally involve discussions with action's staff, beneficiaries, authorities and other stakeholders, discussions on activities under implementation and status of output delivery and preliminary outcomes achieved. These have also the benefit of increasing team building's spirit.

**BOX 36: CHECKLIST OF THINGS TO DO/CONSIDER IN PLANNING, HOLDING AND PROVIDING A FOLLOW-UP OF A MONITORING MEETING**

**Preparation**

Prior to conducting a review meeting, the following tasks should be undertaken by those responsible:

- Confirm who will attend/participate and who will chair the meeting. Specifically consider whether there are opportunities for involving other relevant donors.
- Confirm the date, time and location of the meeting with participants.
- Prepare a draft agenda and distribute it for comment/additions.
- Assemble relevant data/information (including management/monitoring reports) and distribute copies in advance to those attending the review meeting.
- Organise other logistics for the review meeting (e.g. secretarial support, transport, venue, required equipment/materials for presentations, refreshments, etc.).

**The review meeting**

[128] For example, the contractual/financial part is not part of the responsibilities of external monitoring, represented in the Commission practice by the ROM.

[129] According to DEVCO Project Cycle Management Guidelines, "A project 'Governing Body' or 'Steering Committee' is... often required to make strategic decisions on project scope, including required changes in objectives, targets, budget, management arrangements, etc. Such a governing body/committee might therefore meet to review project progress and performance on a periodic basis (i.e. six-monthly or annual), and make the necessary decisions to keep the project 'on track'." When referring to budget support operations, and depending on the specific setting foreseen, these committees, made up of the involved line Ministries, coordinating Ministry, donor partners, other stakeholders, are supposed to review the status of implementation of the reforms targeted, to entertain the required policy dialogue, to analyse the status of achievements of the indicators set out.

Managing the review meeting is primarily the responsibility of the 'chairperson'. The chair should help ensure that:

- The available time is effectively managed, based on the agreed agenda/timetable;
- Each participant is given adequate opportunity to share his/her views (the meeting is not dominated by the loudest/most talkative);
- Key issues are clarified;
- Disagreements are cordially resolved;
- A problem solving approach is taken;
- Agreement is reached (by consensus or vote) on key actions that need to be taken;
- An accurate record of discussions and decisions is taken.

**Follow-up**

Key follow-up actions should include:

- Finalisation and dissemination of a record of key decisions taken/agreements reached.
- Revision to forward work plans as required.

*Source: Strengthening project internal monitoring, Op. Cit.*

- Visits to premises where activities are taking place. At least one ordinary monitoring mission should be foreseen for at least 50% of actions on an annual basis under centralised management; 30% under indirect management.
- On the spot checks: these are more structured action visits, normally involving, beside the activities described above, the formal filling of a checklist and some formal checks[130]. On the spot checks should cover at least 20% of contracts under all management modes.

**BOX 37: GUIDANCE OF THINGS TO DO/CONSIDER BEFORE IMPLEMENTING AN INTERNAL MONITORING MISSION AND/OR AN ON THE SPOT CHECK**

- Collect background documents, including (as appropriate): (i) Action document, (ii) intervention logic diagram, (iii) most recent annual/updated work plan and budget; (iv) previous monitoring/progress report(s); (v) relevant financial statements.
- Familiarise yourself with the content of these documents, and discuss issues with the implementing partner and with other colleagues who may be working on the same area or in similar interventions.
- Clarify the purpose of the visit: What will the visit achieve? Is the purpose of the visit primarily to 'audit/check', or is there also a support/advisory role to be played? What will the implementing agency/stakeholders get out of the visit? **How can you add value**?
- Identify the key issues that need to be addressed during the visit (look at the plan, the key assumptions and any issues raised in previous progress reports). Develop a preliminary list of key questions that it would be useful to ask and have answered.
- Clarify who will/should be involved in the visit, both in terms of the 'monitoring team' and other stakeholders who you wish to meet with. Involve the implementing partner and other donors who may be interested in participating.
- Think through and clarify the proposed approach/methods to be used to collect, record and analyse information: Who do you want to meet, where and when? Do you want to conduct group or individual interviews? Do you want to meet with women separately from men? What do you

[130] For example: presence and utilisation of supplies; verification of presence of stickers mentioning the donor funding; verification of status of implementation of works, whenever relevant; the premises actually built, whether they correspond to the plans in terms of location, characteristics, utilisation, etc., checking of contractual or financial documents, if appropriate.

want to see – physically inspect? What administrative records would you like to see? How will you avoid 'bias' in terms of who you meet and what you are shown by partners/stakeholders who may try to show you only 'success' stories?

- Develop an indicative itinerary for the visit and confirm with those who need to know.
- Clarify the expected output of the visit, including reporting requirements and how information will be 'fed back' to those who need to know.

*Source: Strengthening project internal monitoring, Op. Cit.*

Data collected through the action's monitoring system and related data collection methods is then to be analysed and interpreted[131].

As the Strengthening project internal monitoring guidelines state:

> When thinking about the way in which data should be analysed, different approaches are usually required for quantitative and qualitative data. By definition, quantitative data involves numbers that can be subjected to various forms of statistical analysis. Qualitative data on the other hand usually provides information on people's views, opinions or observations and is often presented (at least initially) in a narrative form (it is also possible to turn qualitative information into a quantitative form). An appropriate balance between the two is often best – with the interpretation of quantitative data being 'enriched' through an understanding of 'what people think'. Also, while some types of project lend themselves to quantifiable monitoring (e.g. those focused on increasing agricultural production, or increasing immunisation coverage) others do not (e.g. institutional capacity building within an environmental research agency).

**Lessons learned should play a key role in data analysis and interpretation (and subsequent reporting).** They are to be considered **as outcomes of a learning process, which involves reflecting upon the experience**. **They are the basis of corrective measures**. The key questions to be answered are:

- What has and has not worked?
- Why? Which were the enabling and limiting factors?
- How will these lessons be considered? Which stakeholder(s) will act upon them?

**BOX 38: EXAMPLES OF SUPPORT QUESTIONS IN ASSESSING DATA ANALYSIS AND USE[132]**

The following represents a tentative list of questions one should address:

- Is information being analysed at an operational level to help implementers understand what they are doing before being passed up to higher levels?
- Is the nature of the analysis appropriate and useful? (e.g. are comparisons made between what was planned and actual outcomes)?
- Is there a functioning review system for bringing together action stakeholders to make decisions based on the available information?
- How does this operate and who is involved? Is it coordinated with other donors?

*Source: DG NEAR*

[131] As mentioned under Step 7 – Design the data collection systems, and under Table 3: Monitoring actors and relevant key responsibilities, this task is also to be carefully resourced in financial and human terms.

[132] All organisations collect, analyse and distribute information, even if they do not call it 'monitoring'. *Strengthening project internal monitoring Op. Cit.*

## 11.3. Internal monitoring reporting

### *11.3.1. Focus of reporting*

**The utility of a monitoring exercise also resides in the quality of reporting. And, the quality of internal monitoring reporting has a direct effect on**:

- The quality of ROM,
- The quality of results' reporting done by ROM contractors. And this is even more important considering that this is the basis of corporate reporting whenever applicable.
- The quality of EAMR, and
- The quality of relevant action steering Committees and relevant management committee under IPA II.

#### 11.3.1.1. Implementing partner reporting

Internal monitoring reporting, and especially that of implementing partners, needs to go beyond activities and input use reporting (and this is unfortunately one of the common concerns in internal monitoring). Indeed, as mentioned elsewhere, monitoring is embedded into an action's cycle and is expected to answer to the following question: **What is the action doing and how much implementation is progressing**? In order to answer this question, **both the implementing partners and EC HQs/EUD programme managers need to concentrate on providing a final value of measurement to the indicators' targets defined for Action Documents at the activity, output and outcome level**.[133] Activity reporting is therefore only one component and not the most important one since it is at the lowest level of the results' chain, of monitoring reporting. As stated in the internal monitoring guidelines prepared by DG EuropeAid in 2007: the critical issue is that those implementing the action understand that there is a results hierarchy, and that their activities therefore have a higher 'purpose' which must, in time, be demonstrated. [134]

**In order to avoid failure in achievement monitoring's key objectives (assess the progress towards objectives and identify implementation problems (and devise relevant corrective measures)), EC HQs and EUD programme managers must include in the implementing partner's ToR the need to establish this kind of sound reporting.**

TABLE 14: SUGGESTED CONTENT OF MAIN TYPES OF REPORTS providing information on key elements to be considered in internal monitoring note/reports[135].

**TABLE 14: SUGGESTED CONTENT OF MAIN TYPES OF REPORTS**

| Inception Report (First Annual Plan) | Progress Report and Annual Plan | Completion Report |
|---|---|---|
| **Table of contents and list of abbreviations** | **Table of contents and list of abbreviations** | **Table of contents and list of abbreviations** |
| **1. Introduction** | **1. Introduction** | **1. Introduction** |
| 1 page that summarises (i) basic action data (name, location, duration, value, key stakeholders, purpose and key results, etc.) (ii) the status of the action at the | 1 page that summarises (i) basic action data (name, location, duration, value, key stakeholders, purpose and key results, etc.) (ii) the status of the action at the | 1 page that summarises (i) basic action data (name, location, duration, value, key stakeholders, purpose and key results, etc.) (ii) the status of the action at the |

[133] This last level will of course depend on the implementation status.

[134] *Strengthening project internal monitoring Op. Cit.*.

[135] In the case of reporting done by national authorities and other donors, the format should be jointly discussed.

| Inception Report (First Annual Plan) | Progress Report and Annual Plan | Completion Report |
|---|---|---|
| time of reporting; and (iii) who has prepared the report, why and how. | time of reporting; and (iii) who has prepared the report, why and how. | time of reporting; and (iii) who has prepared the report, why and how. |
| **2. Executive summary and recommendations**<br>Concise summary (i.e. 2 pages) of the main issues and recommendations for the attention of key decision makers. | **2. Executive summary and recommendations**<br>Concise summary (i.e. 2 pages) of the main issues and recommendations for the attention of key decision makers. | **2. Executive summary and recommendations**<br>Concise summary (i.e. 2 pages) of the main issues and recommendations for the attention of key decision makers. |
| **3. Review of action design (relevance, feasibility and any changes required to design)** (up to 10 pages)<br>3.1 Policy and action opportunity framework (enabling and limiting factors), including linkage to other on-going operations/activities<br>3.2 Objectives to be achieved (Overall Objective, specific objectives, outputs)<br>3.3 Activities<br>3.4 Resources (political, technical, financial and human inputs) and budget<br>3.5 Management, coordination and financing arrangements<br>3.6 Monitoring, review and evaluation arrangements<br>3.7 Key Risk Management/Sustainability issues (update) | **3. Review of Progress and Performance to date (comparing against plan – efficiency and effectiveness)** (up to 10 pages)<br>3.1 Policy and action opportunity framework (enabling and limiting factors - including on institutional capacity)), including linkage to other on-going operations/activities<br>3.2 Progress towards achieving objectives (Objective, specific objectives, outputs)<br>→ Analysis of results achieved based on reporting against planning/programming (or updated) indicators<br>→ Problems/limitations encountered<br>3.3 Activities undertaken<br>3.4 Resources (political, technical, financial and human inputs) and budget used<br>3.5 Management, coordination and financing arrangements<br>3.6 Overall conclusion: Key issues arising and action required[136] | **3. Review of Progress and Performance at completion (comparing against plan – efficiency, effectiveness and impact)** (up to 10 pages)<br>3.1 Policy and action opportunity framework (enabling and limiting factors), including linkage to other on-going operations/activities<br>3.2 Objectives achieved (Overall Objective, specific objectives, outputs)<br>→ Analysis of results achieved based on reporting against planning/programming (or final) indicators<br>→ Problems/limitations encountered<br>3.3 Activities undertaken<br>3.4 Resources (political, technical, financial and human inputs) and budget used<br>3.5 Management, coordination and financing arrangements<br>3.6 Sustainability issues |
| **4. Work plan for the next period (Annual Plan)**<br>4.1 Results to be delivered – quantity, quality and time<br>4.2 Activity schedule – including any key milestones and lead responsibilities<br>4.3 Resource schedule and budget<br>4.4 Risk management & sustainability | **4. Recommendations for the remaining phase(s)**<br>If relevant, organised by topics/clusters. | **4. Lessons learned**[137]<br>4.1 Policy and opportunity framework (enabling and limiting factors - including on institutional capacity)<br>4.2 Process of action planning/design<br>4.3 Action scope (objectives, resources, budget, etc.)<br>4.4 Action management/coordination arrangements and stakeholder participation<br>4.5 Action financing arrangements<br>4.6 Risk management and Sustainability |
| | **5. Work plan for the next period (Annual Plan)** | **5. Recommendations for a next phase of the same action/similar action (if** |

[136] The report must describe how evidence from lessons learned and evaluation (if available) are being used to inform any adjustments needed.

[137] Lessons learned, including evidence from evaluations, if available, must be clearly and specifically described.

| Inception Report (First Annual Plan) | Progress Report and Annual Plan | Completion Report |
|---|---|---|
| | 5.1 Results to be delivered – quantity, quality and time<br>5.2 Activity schedule – including any key milestones and lead responsibilities<br>5.3 Resource schedule and budget<br>5.4 Risk management and sustainability | relevant) |
| **Annexes**<br>• Updated Log frame Matrix<br>• Monitoring and Evaluation Plan, including revised overall targets<br>• Updated Annual Work plan for first year<br>• Updated Annual Resource Schedule and budget<br>• Other | **Annexes to the Annual Plan**<br>• Updated Log frame Matrix<br>• Summary performance data (results, milestones and expenditure – for reporting year and cumulative to date)<br>• Updated Annual Work plan for next period<br>• Updated Annual Resource Schedule and budget for next period<br>• Other | **Annexes**<br>• Summary performance data (purpose, results and expenditure – cumulative to date)<br>• Other |

*Source: DG NEAR, adapted from Strengthening project internal monitoring, Op. Cit.*

#### 11.3.1.1. EC HQs/EUD reporting

This type of reporting is linked to the internal monitoring missions and on the spot checks. It is done on the basis of checklists. These should be filled in following the missions carried out based on a set of questions/issues to be investigated during the visit (refer to BOX 37: GUIDANCE OF THINGS TO DO/CONSIDER BEFORE IMPLEMENTING AN INTERNAL MONITORING MISSION AND/OR AN ON THE SPOT CHECK). Specific questions might address the different action components and contracts involved (service, work, supply, grant)[138].

Further to the internal monitoring missions and on the spot checks, a monitoring note/report should be prepared and submitted to the Head of Operation/Unit and, whenever relevant, to the relevant Director for their perusal and potential action.

Main findings and recommendations and actions taken on past monitoring are reported, as appropriate, in the EAMR and discussed in the relevant monitoring platforms. The fact of having a standardized template (to which additional elements can be added thought depending on the specificities of the intervention or on the specific needs) would contribute facilitating corporate reporting.

[138] Annexes 11 to 15 provide checklists' templates.

**BOX 39: POTENTIAL BENEFITS OF EC HQS/EUDS STAFF USING QUESTION CHECKLISTS**

- They help to ensure that key issues are covered during field monitoring visits;
- They help to ensure consistency and comparability of reporting;
- The discipline of checklists helps to institutionalise a system of action monitoring which assists incoming staff to familiarise themselves with the action and thus become effective more quickly;
- A structured information collection and reporting system is an important ingredient for developing an institutional memory. Ad hoc systems tend to become very personalised and break down when the officers leave; and
- The completed checklists may provide some raw data for subsequent analysis, if the questions are adequately structured.

*Source: Strengthening project internal monitoring, Op. Cit.*

### *11.3.2. Internal monitoring note/reports vs other information sharing tools*

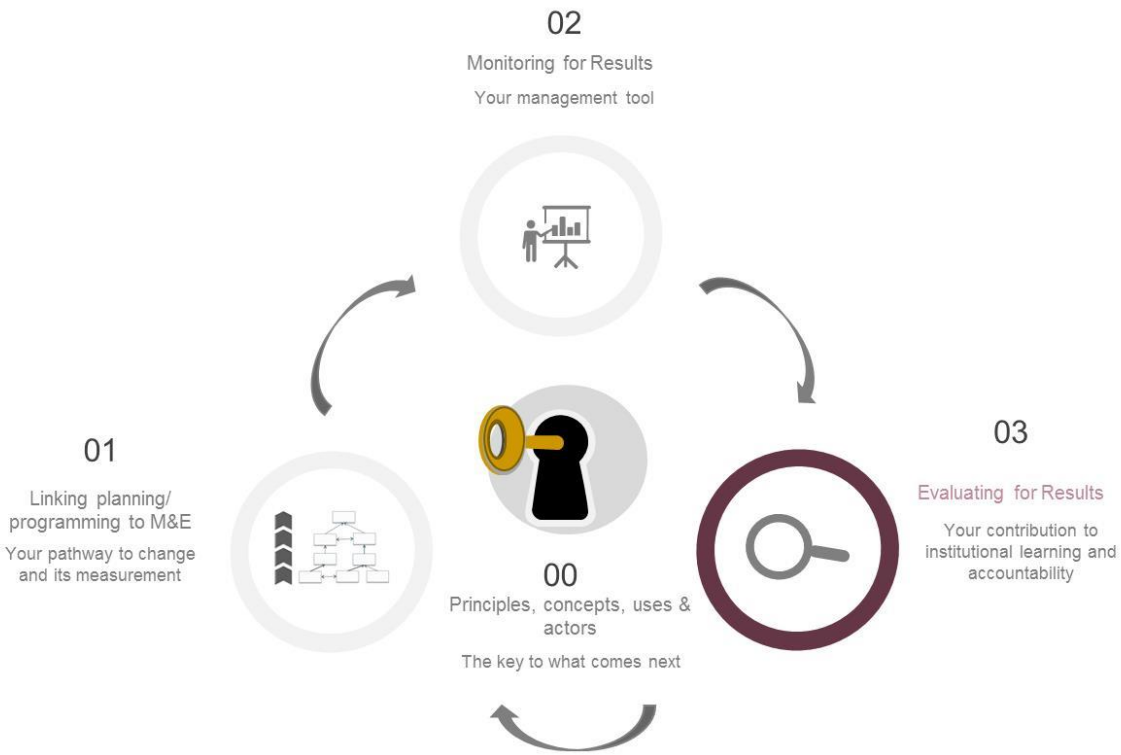
Internal monitoring note/reports are not meant to replace other information sharing and dissemination tools, nor other types of monitoring activities, such as formal and informal meetings with key stakeholders directly and/or indirectly involved in the action for example. As such, action reports should not try to provide full discussion of all issues or concerns, they need to be kept clear and concise. To the extent possible, they should only contain high value information that needs to go on record. In this regard, even though it can be advised to report at least twice a year, reporting should be appropriate to need. [139]

## 11.1. Dissemination and Follow-up

Other than stakeholders directly involved in monitoring (and as such in action's management level), it is also expected to disseminate monitoring results to actors not directly involved in the exercise but having, for some of them, a key role in the follow-up as well:

- At the EC/EUD level, this would be primarily the case of the Head of Unit/Operations that would also be involved in the follow-up, and other programme managers working in the same sector and/or in similar interventions within the same Unit/EUD. The director would be kept informed via the EAMR, unless specific gravity might necessitate sharing the monitoring note/report with him/her.
- Nevertheless, keeping the institutional learning purpose in mind, it is also highly recommended that monitoring results (and implementing partners' completion reports) that are considered of good quality are shared to a wider audience via DG NEAR HQ's Monitoring & Evaluation service (that would be acting as a repository of monitoring practices and lessons learned).
- At the national authorities' level:
- To the national aid coordinator and relevant action steering committees in both ENI and IPA II
- In the case of IPA II, monitoring results to be shared with the National IPA Coordinator (NIPAC), via IPA monitoring committee, and with Sectorial monitoring committees. These would have a role in the follow-up.
- Other stakeholders, mainly at the level of beneficiaries.

[139] *Strengthening project internal monitoring Op. Cit..*




# PART 3: EVALUATING FOR RESULTS

**BOX 40: SUMMARY OF KEY MESSAGES ADDRESSED IN PART 3**

***The best guarantee for the quality and usefulness of an evaluation is good management.***

***Full ownership of an evaluation by the service launching it increases the likelihood that all EC HQs/EUDs use evaluation findings in internal self-reflection, future interventions, and in the dialogue within the European Commission, with other European Institutions, and with partner countries.***

***The Commission Belter Regulation guidelines introduces different mandatory steps:***

***An Inter-service steering group (ISG) is to be set up for each evaluation managed by DG NEAR HQs. For the other evaluations, a Reference Group (RG) is to be formed.***

***The publication of a Roadmap is mandatory only for strategic evaluations***

***A consultation strategy for all evaluations and an Open Public Consultation only for major strategic evaluations***

***An Evaluation Staff Working Document, only for major strategic evaluations***

***Reconstructing the Intervention logic and identifying evaluation questions are a shared responsibility of evaluation managers (as steering actor of the Interservice group/Reference group) and of evaluators.***

*Source: DG NEAR*

## 12. TYPES

There are different types of evaluations[140], according to different types of criteria.

### 12.1. By purpose

Evaluations may be **formative**, when they are intended to improve performance, most often conducted during the implementation phase of interventions. They provide assessment and lessons learned, on the basis of the past interventions, drawbacks and good practices identified. Formative evaluations may also be conducted for other reasons such as compliance, legal requirements or as part of a larger evaluation initiative. Ex-ante and mid-term evaluations are formative in nature.

They can be **summative**, when they are conducted at the end of an intervention (or a phase of that intervention) to determine the extent to which anticipated outcomes were produced. Summative evaluation is intended to provide information about the worth of the intervention. They provide a forward-looking

[140] Definitions are directly taken/adapted from OECD-DAC, 2010a.

assessment and recommendations for actions, policies which will be prepared in the near further, based on the lessons learned from the past interventions, underpinned by the monitoring and evaluation results. Mid-term and ex-post evaluations are summative.

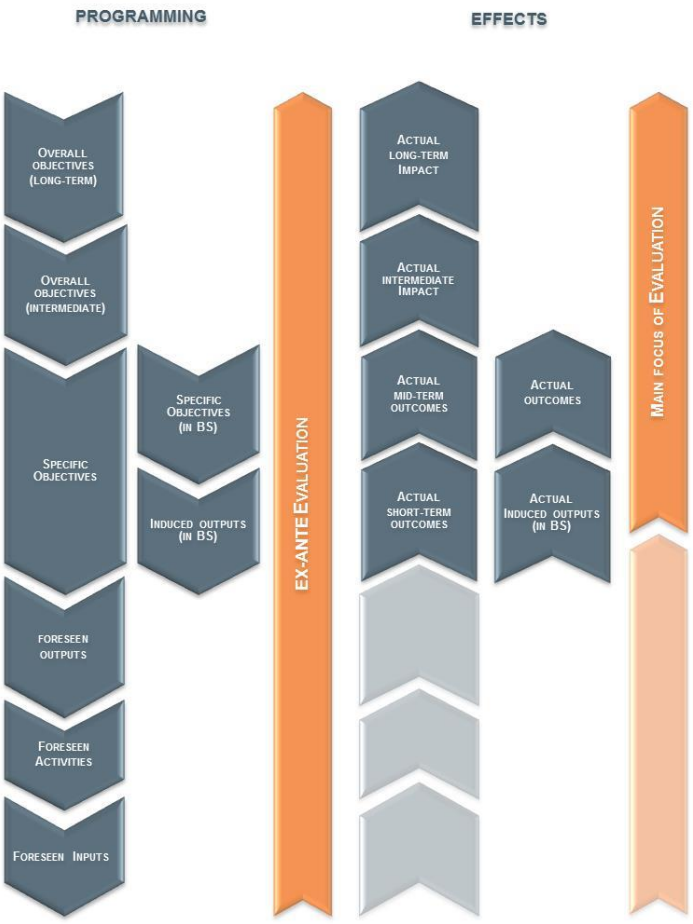
A good practice in the sector/thematic/ country bases evaluations is to encompass both forward and backward looking perspectives in the design of one evaluation.

A **meta-evaluation** [141] & [142] evaluates the evaluation, both the process and the deliverables. It is time-consuming and can be intrusive—the meta-evaluator needs to have full access to the entire process. It differs from quality control because it implies original research and is best suited to situations in which innovations are tested. Da facto, meta-evaluations are sometimes assimilated to a synthesis evaluation, which is a different exercise.

A **synthesis evaluation** is designed to aggregate findings from a series of evaluations

A **cluster evaluation** is that of a set of related activities, projects and/or programs.

**FIGURE 14 : EVALUATION, ITS PLACE WITHIN THE INTERVENTION LOGIC**




*Source: DG NEAR*

### 12.2. By agent

**Self-evaluation**[143]: an evaluation by those who are entrusted with the design and delivery of an intervention.

**Internal evaluation:** Evaluation conducted by a unit and/or individuals reporting to the management of the donor, partner country, or implementing organization. They do not involve an external contractor.

**External evaluation:** The evaluation conducted by entities and/or individuals outside the donor and implementing organizations. They are contracted by the donor and/or implementing organizations services.

[141] The term meta-evaluation, or meta-analysis, is also used to indicate reviews which aggregate information coming from different evaluations addressing the same topic, possibly supported by complementary findings.

[142] They involve appointing senior evaluator(s) to evaluate the evaluation, both the process and the deliverables. It is time-consuming and can be intrusive—the meta-evaluator needs to have full access to the entire process. It differs from quality control because it implies original research and is best suited to situations in which innovations are tested.

[143] Self-evaluations possess great value within a learning evaluation: they spur self-reflectiveness. The coincidence of interests between the evaluator and the programme manager of the intervention ensures ownership. However, self-evaluations only rarely question organisational practices and strategies and the premises for the intervention design and rationale—usually, organisation members are deeply steeped in their organisation main discourse and practice. In addition, they are usually considered as being less credible vis-à-vis external stakeholders and the general public. This provided a judgement.

**Independent evaluation:** An evaluation carried out by entities and persons free of the control of those responsible for the design and implementation of the intervention. They are under the responsibility of for example independent evaluation units, responding directly to boards or elective assemblies.

Note: The credibility of an evaluation depends in part on how independently it has been carried out. Independence implies freedom from political influence and organizational pressure. It is characterized by full access to information and by full autonomy in carrying out investigations and reporting findings.

**Mixed evaluation:** They combine elements of internal and external evaluations. The work is carried out by the donor and implementing organizations services, as well as by an external contractor/evaluator.

**Joint evaluation[144]:** An evaluation, in whole or in part, being carried out by more than one organization, and/or evaluating the interventions being implemented by more than one organization[145]. The definitions vary in terms of the scope of interaction or "jointness" and the nature and mechanism of exchange. The joint evaluation could take form of:

- Joint data collection and/or or exchanging assessments-with external actors where each partner conducts its own analyses and separate report
- Collaborative evaluation-each partner is mutually and equally responsible for the evaluation design, implementation and development of joint recommendations.

Joint evaluations can help overcome attribution problems in assessing the effectiveness of interventions and strategies, the complementarity of efforts supported by different partners, the quality of aid coordination, etc. It is necessary to consider this option of organising the evaluation when appropriate conditions are available.

### 12.3. By timing in the intervention's cycle

**Impact assessment** is defined by the Better Regulation as an exercise about gathering and analysing evidence to support policy making. In this process, it verifies the existence of a problem, identifies its underlying causes, assesses whether EU action is needed, and analyses the advantages and disadvantages of available solutions.

**Ex-ante evaluation**, performed before the implementation of an intervention. Under the Better Regulation, they are very similar to an impact assessment but they mostly relate to project/programme evaluations.

**Mid-term evaluation**: Evaluation performed towards the middle of the period of implementation of the intervention. They encompass both forward and backward looking perspectives.

As shown in the **Figure 15: The planning/programming and evaluation cycle**, mid-term evaluations arrive too early in intervention's implementation for outcomes to have materialised yet and to guide mid-term reviews or to solve implementation riddles. **Programme managers** can get relevant knowledge from:

[144] Joint evaluations are well known in the development and humanitarian sector and are still relatively new for DG NEAR, though a joint evaluation on the experience of budget support in Morocco has already been finalised in 2014. It can be found at: https://ec.europa.eu/europeaid/joint-strategic-evaluation-budget-support-operations-morocco-2005-2012_en.

According to the DAC-Network Development Evaluation Inventory (2014) only 16% of submitted evaluations were joint evaluations. The inconveniences in their planning is that they are generally more costly; planning processes are time consuming, as are negotiations on scope, methodology and evaluation questions and timing of needed formative feedback becomes more challenging with more parties involved. The benefits that they bring are numerous: mutual capacity development; harmonization and reduced transaction costs; increased participation of developing countries; increased objectivity, transparency, and legitimacy of the evaluation and its impact; ability to address a broader scope of evaluation question.

[145] Beck and Buchanan-Smith (2008).

- ex post evaluation of similar interventions
- impact evaluations of early pilots (if they were implemented) and impact evaluations of previous relevant interventions
- ex ante evaluations, which highlight and analyse alternative options for the intervention and make the intervention logic explicit, thus allowing the programme manager to compare what is really happening with the hypothesised sequences.

**Ex-post evaluation**: Evaluation of an intervention after it has been completed.

Note: It may be undertaken directly after or long after completion. The intention is to identify the factors of success or failure, to assess the sustainability of outcomes and impacts, and to draw conclusions that may inform other interventions.

DG NEAR interventions (both annual and multiannual)[146] overlap. The preparation of an intervention starts well before the end of the preceding one. It is, therefore, not possible to use the ex post evaluation of an intervention in the preparation of the subsequent one. P**rogramme managers engaged in intervention preparation** should, therefore, in addition to asking, when relevant and feasible, for an ex ante evaluation:

- Use knowledge from the mid-term evaluation and/or the ROM report for the previous period
- Gather information from the monitoring system of the previous period
- Use ex post evaluations of previous interventions in the country/sector of interest
- Ask for an ex-ante evaluation.

### *12.3.1. Better regulation requirements on the timing of evaluations*

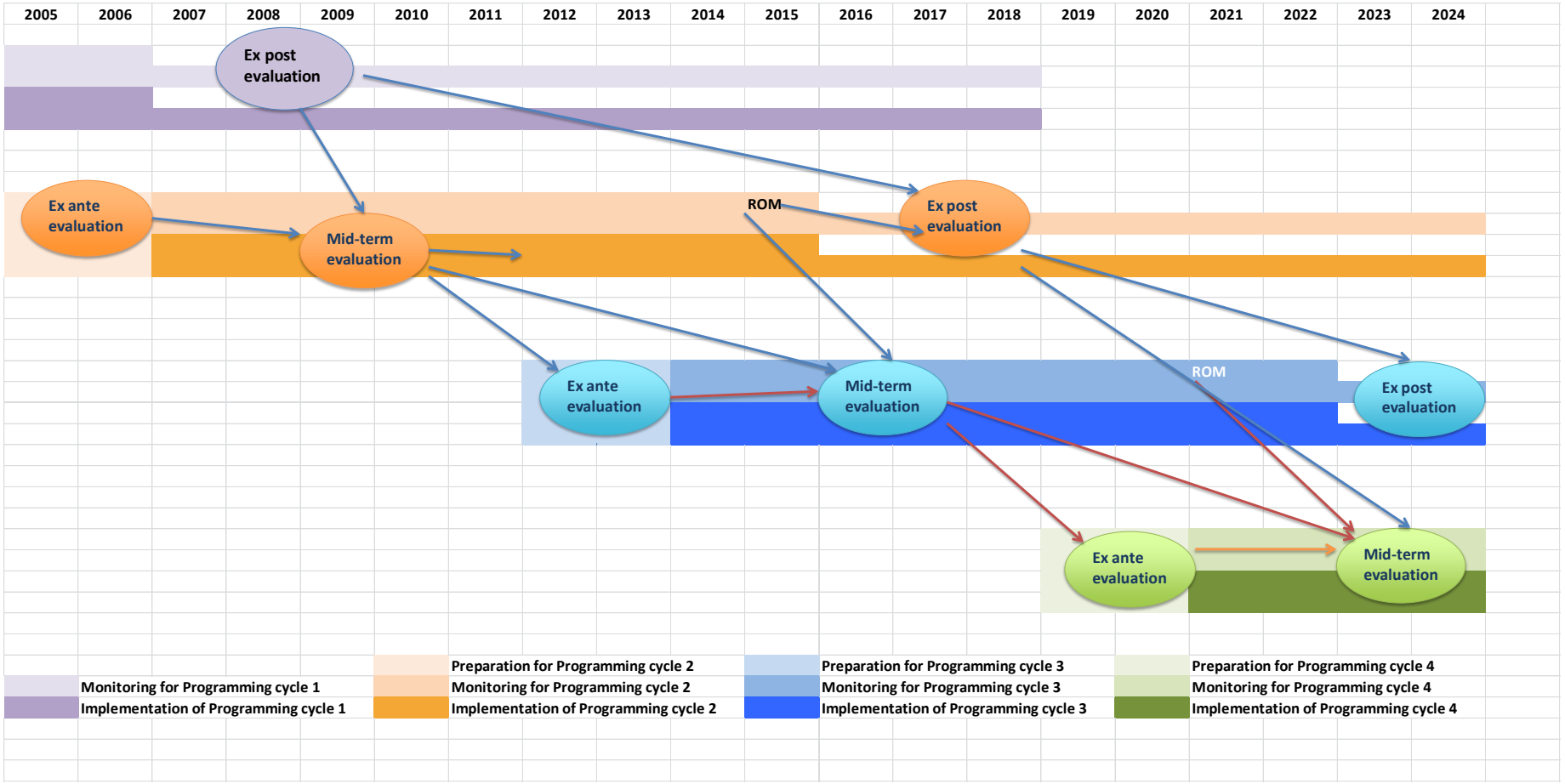
The "**evaluate first**" principle requires that every proposal for a new intervention or for amending an intervention be accompanied by an evaluation of pre-existing interventions or regulations. Evaluations must come before Impact Assessments. The principle therefore, closes the evaluation cycle and completes the **planning/programming-evaluation cycl**e, giving a new ordering to the sequence of **ex post evaluations**, **ex ante evaluations**, **mid-term evaluations** and **ongoing evaluations**, and again **ex post evaluations**.

Launch and promptly evaluate a pilot, when an intervention includes a very innovative activity for which there is little evidence whether it will work in a particular environment. The pilot should be implemented and evaluated at the end of the previous planning/programming period.[147]

[146] For a description of the way the programming works for ENI and IPA II, please refer to: https://myintracomm.ec.europa.eu/dg/near/whatwedo/neighbourhood/Pages/programming.aspx, for ENI, and https://myintracomm.ec.europa.eu/dg/near/whatwedo/financial-assistance/Pages/ipa-programming-implementation.aspx, for IPA II.

[147] This might not be feasible for all sectors: evaluations results for a pilot might not be valid for larger scale actions, because scale matters (e.g., for activities such as campaigns against rabies or transportation network improvements) or outcomes need time to materialise (e.g., for actions in favour of small farmers).

**FIGURE 15: THE PLANNING/PROGRAMMING AND EVALUATION CYCLE**




*Source: DG NEAR*

The Better Regulation guidelines clearly state that the regulation requirements on evaluation must be intended as minimum requirements: both at Headquarters, in Delegations, and in partner countries, it is recommended to conduct more evaluation than the ones strictly mandated (refer to TABLE 15: COMPULSORY VS. RECOMMENDED EVALUATIONS IN THE PLANNING/PROGRAMMING CYCLE)

**TABLE 15: COMPULSORY VS. RECOMMENDED EVALUATIONS IN THE PLANNING/PROGRAMMING CYCLE**

| Actors | Ex-ante evaluations | | Interim/Mid-term evaluations | | Final evaluation | Ex-post evaluations | |
|---|---|---|---|---|---|---|---|
| | At instrument level (strategic)/ Impact assessment | At action level | At instrument level (strategic) | At action level | At action level | At instrument level (strategic) | At action level |
| **DG NEAR HQ's monitoring & evaluation service** | C | R | C (IPA II)<br>R (ENI) | | | C | R |
| **DG NEAR – operational units** | | R | | | R | | |
| **EUDs** | | R | | C[148] (in direct management for actions above EUR 5 M)<br><br>R (where relevant and useful) actions less than EUR 5M) | R (in direct management for action above EUR 5 M and/or innovative actions less than EUR 5M) | | C (in direct management for actions above EUR 5 M)[149]<br><br>R (where relevant and useful) |
| **IPA II/ENI beneficiaries** | | R | | C (in direct management for actions above EUR 5 M)<br>R (where relevant and useful)<br><br>C (indirect management for action above EUR 5 M and/or innovative actions less than EUR 5M) | C (indirect management for action above EUR 5 M) | | C (in direct management for actions above EUR 5 M) [150]<br><br>R (where relevant and useful) |

*Source: DG NEAR*

Legend: C- compulsory; R- recommended.

### 12.4. By scope[151]

**Project[152] evaluation:** Evaluation of an individual action designed to achieve specific objectives within specified resources and implementation schedules, often within the framework of a broader action.

[148] Indeed, Art. 18 of the Commission Delegated Regulation (EU) No 1268/2012 of 29 October 2012 on the rules of application of Regulation No 966/2012 " All actions, including pilot projects and preparatory actions, where the resources mobilised exceed EUR 5 000 000 shall be the subject of an interim and/or ex post evaluation in terms of the human and financial resources allocated and the results obtained in order to verify that they were consistent with the objectives set". In the practice of DG NEAR, it is proposed to increase this amount to €10 million. This threshold might disappear in the framework of the new Financial Regulation.

[149] Not compulsory if a mid-term evaluation has taken place.

[150] Not compulsory if a mid-term evaluation has taken place.

[151] Often, these evaluations are linked: thematic evaluations, for example, take into consideration actions' evaluations.

**BOX 41: PROJECT EVALUATION REQUIREMENTS, AS PER THE BETTER REGULATION AND THE FINANCIAL REGULATION [153] AND ITS RULES FOR APPLICATION[154]**

The **Better Regulation** guidelines allow for simpler requirements for project evaluations: it is not necessary to apply all the requirements (e.g., drafting roadmaps, conduct full-fledged public consultations, prepare a Staff Working Document, or consider all 5 key evaluation criteria). For "individual actions, groups of actions or sub-activities where their findings will feed into an overarching evaluation. This is particularly relevant for external actions where findings coming from evaluations of country interventions, specific delivery methods/tools or elements of certain themes feed into major evaluations including of legal instruments".

There is also a requirement, stipulated by the **Financial Regulation** and its **Rules for Application**, requiring that all interventions or activities, including pilot projects and preparatory actions, where the resources mobilised exceed EUR 5 000 000 shall be the object of a mid-term and/or ex post evaluation in terms of the human and financial resources allocated and the results obtained in order to verify that they were consistent with the objectives set. In this context, an evaluation on projects above EUR 5 M is compulsory. Nonetheless, DG NEAR suggests evaluations to take place for projects above €5 million, though innovative or particularly sensitive projects are also suggested to be evaluated[155].

*Source: Better Regulation*

Note: Cost benefit analysis is a major instrument of project evaluation for projects with measurable benefits, for impact assessment/ex-ante evaluations. When benefits cannot be quantified, cost effectiveness is a suitable approach.

**Programme[156] evaluation:** Evaluation of a set of actions, marshalled to attain specific global, regional, country, or sector objectives.

**Country[157]/Regional Evaluation**: Evaluation of one or more donor's or agency's portfolio of interventions, and the assistance strategy behind them, in a partner country or in a region.

The purpose of the evaluation is the entire EU strategy for a partner country. These strategies include many forms of aid delivery, policy dialogue, political conditionalities, links with trade policy, and interventions in different fields. They include concluded as well as on-going interventions. The evaluations assess the changes that the strategy has contributed to in each country, for example how the strategies implemented in the countries reduced its gap vis-à-vis the *acquis communautaire*. They take into account different interventions implemented in different sectors, aiming at assessing results achieved in a given country in the perspective to achieve the strategic objectives set out in the relevant planning document (Indicative Strategy Paper/Single Support Framework, etc.) using the different instruments available. Country evaluation can substantially benefit from intervention evaluations that have been previously performed.

[152] A project is defined as a series of activities aimed at bringing about clearly specified objectives within a defined time-period and with a defined budget. (from the EC PCM guidelines – 2004).

[153] Regulation (EU, Euratom) No 966/2012 of the European Parliament and of the Council of 25 October 2012 on the financial rules applicable to the general budget of the Union and repealing Council Regulation (EC, Euratom) No 1605/2002.

[154] Commission Delegated Regulation (EU) No 1268/2012 of 29 October 2012 on the rules of application of Regulation (EU, Euratom) No 966/2012 of the European Parliament and of the Council on the financial rules applicable to the general budget of the Union.

[155] See ARES Note n. 4080381.

[156] A programme may cut across sectors, themes and/or geographic areas.

[157] An example of a country evaluation can be found at: http://ec.europa.eu/europeaid/sites/devco/files/strategic-evaluation-cooperation-ec-jordan-1340-main-report-201502_en.pdf.

**This type of evaluation is considered of a strategic nature in DG NEAR.**

**Sector evaluation**[158]: Evaluation of a cluster of interventions in a sector or sub-sector within one country or across countries, all of which contribute to the achievement of a specific goal.

The interventions can support each other, share the same features in different parts of the country or in the same area, or be in a time sequence. It may be very difficult to combine data across interventions to accurately measure combined outcomes. The strong point of these evaluations reside in the possibility of comparing different institutional arrangements, comparing how similar interventions act in different contexts, and in analysing complementarities or conflicts and synergies or, on the contrary, crowd-out effects or perverse effects, shedding light on the positive and negative aspects which single intervention evaluations and monitoring may overlook.

**Thematic evaluation**[159]**:** Evaluation of a selection of interventions, concluded or still on going, all of which address a specific policy priority that cuts across sectors, countries or regions. The theme may correspond to an expected impact (e.g. competitiveness of SMEs) or to a field of interventions (e.g. R&D)[160]. It can also correspond to a crosscutting issue (e.g., gender equality). The aim is to address the link between results and the existing sector strategies set out. **This type of evaluation is considered of a strategic nature in DG NEAR.** When they address key pillars of the SG's strategy, like the **fundamentals, they need to be fully compliant with the Better Regulation.**

**Other strategic evaluations** under the responsibility of DG NEAR HQ's Monitoring & Evaluation service, and that do not specifically address neither a country/region nor a sector or theme, are the following:

- **Aid modality evaluation**[161]**:** They look at the ways specific aid delivery methods (like budget support[162], or twinnings, grants, service, supply, and work contracts) take place. These evaluations encompass programmes, instruments, and aid modalities across countries and over long time spans, including concluded and on-going interventions. Therefore, they combine case studies and the assessment of the strategy or of the instrument as a whole.
- **Financial instruments (like ENI or IPA)**[163]**:** these evaluations look at the financial instrument as defined at regulation level, being typically designed to implement a specific policy. **It is to be fully compliant with the Better Regulation.**
- **Policy evaluations**, aiming at apprehend, looking at the different modalities by which they are implemented (policy dialogue, spending actions, etc.) the capacity to achieve the set

[158] An example of sector evaluation can be found at: http://ec.europa.eu/enlargement/pdf/financial_assistance/phare/evaluation/2015/turkey_osh_eval_final_report.pdf.

[159] An example of a thematic evaluation, on Roma, can be found at: http://ec.europa.eu/enlargement/news_corner/key-documents/index_en.htm?key_document=08012624887bedda.

[160] European Commission, 2013, Glossary

[161] An example of a budget support evaluation can be found at: https://ec.europa.eu/europeaid/joint-strategic-evaluation-budget-support-operations-morocco-2005-2012_en.

[162] To respond to the specific challenges posed by the evaluation of budget support, the OECD DAC Evaluation Network has developed a methodology for assessing the chain of results and induced impacts of support at country level. The approach systematizes choices and activities and provides suggestions on the choice of methods and techniques. The evaluation methodology can be found at: http://www.oecd.org/dac/evaluation/dcdndep/Methodological%20approach%20BS%20evaluations%20Sept%202012%20_with%20cover%20Thi.pdf. The methodology was tested in evaluations of budget support in Tunisia, Mali and Zambia in 2010/2011, and, after revision, used – among others – in the Joint evaluations carried out in Morocco (2014) and Uganda (joint IEG and EU team of evaluators 2015).

[163] An example of a country evaluation can be found at:: http://ec.europa.eu/enlargement/pdf/financial_assistance/phare/evaluation/2013/ipa_interim_meta_evaluation_report.pdf.

policy objectives. When they address the main policy of the DG, they need **to be fully compliant with the Better Regulation.**

- **Fitness check:** As per the Better Regulation, a Fitness Check examines available evidence to judge the cumulative results of a group of measures of different nature (spending actions, regulations, etc.) that share a relationship (e.g., common objectives). Fitness Checks focus on synergies and inefficiencies among the group of measures. **It is to be fully compliant with the Better Regulation.**

Each type presents specific challenges and falls within the responsibility of different actors (refer to **BOX 42: SERVICE LAUNCHING THE EVALUATION, BY SCOPE OF EVALUATION**).

**BOX 42: SERVICE LAUNCHING THE EVALUATION, BY SCOPE OF EVALUATION**

Depending of the evaluation's scope, the service that launches the evaluation differs:

- DG NEAR's Evaluation service launches and manages strategic evaluations (country/region[164], thematic, policy, aid modality, instrument). They are launched via the COM 2015 framework contract (only available for strategic evaluations). This framework contract provides for the Evaluation Methodology for European Commission External Assistance, developed by DEVCO, to be used. This applies of course to evaluations launched by DG NEAR via this framework contract.
- DG NEAR HQ's operational Units launch and manage evaluations of centrally managed actions (project, programme). They are launched via the Benef framework contract.
- EUDs launch and manage evaluations of decentralised managed actions (project, programme) and of sector support. They are launched via the Benef framework contract.
- ENI/IPA II beneficiaries, under full-decentralised management, launch and manage, sector and actions (project, programme) evaluations using the BENEF FWC.

*Source: DG NEAR*

### 12.1. Other

**Process evaluation** An evaluation of the internal dynamics of implementing organizations, their policy instruments, their service delivery mechanisms, their management practices, and the linkages among these.

**Capacity development evaluation**[165]. An evaluation of capacity development and/or strengthening at the individual and/or organisational level.

## 13. EVALUATION MANAGEMENT PHASES: PLANNING, IMPLEMENTATION, REPORTING, FOLLOW-UP

This section deals with the phases of evaluation management—from the decision to launch an evaluation to dissemination and follow-up. **TABLE 16: THE MAIN PHASES IN DG NEAR EVALUATIONS WITH KEY MOMENTS AND DELIVERABLES, BY EVALUATION TYPE** shows which steps are relevant for each type of evaluation.

Although **this section focuses on external (both strategic and project/programme) evaluations**, the main steps and the respective responsibilities of the evaluation manager and of the evaluators' teams are

[164] The exception is represented by countries in full-decentralised management, where the responsibility of the management of financial cooperation (both contracting and paying) is delegated to the national authorities.

[165] A Rapid Assessment of Capacity Development (RAC) was developed on behalf of the EC by a team led by Enzo Caputo. The RAC aims to assess the impact that the interaction with an intervention generates at individual and/or organisational level in terms of capacity development and/or strengthening. It is not an evaluation of the intervention, nor the evaluation of the technical assistance component of an intervention. The evaluation methodology (and standard ToR) can be found at: http://capacity4dev.ec.europa.eu/public-cd-tc/minisite/rac.

valid for internal evaluations as well. Also, while strategic evaluations have lengthier phases and additional deliverables (the desk report is for example not usually requested in project/programme evaluations), this does not preclude the evaluation manager to ask evaluators to do the same type of analysis, even if in a less structured form, in each phase. For example, the identification of hypotheses to be tested in the field, that emerge from a desk analysis, would need to be presented to the evaluation manager and the Inter-service Steering group/Reference group before starting the field phase, even if they are not to be presented in a specific deliverable.

Throughout the evaluation process, evaluation managers will have to use the EVAL module[166]. This is now a mandatory requirement. Its utilisation starts with the setting up of the Inter-service Steering group/Reference group and the preparation of the terms of reference.

**The quality of an evaluation resides in both the process and its products**. As mentioned earlier, the best guarantee for the quality and utility of an evaluation is good management: protecting quality and credibility of an evaluation is primarily the evaluation manager's responsibility.

**Evaluation management is key in ensuring quality and utility of an evaluation**. The way it is performed determines to what extent its commissioners (the services having launched the evaluations) own the evaluation. Ownership, in turn, increases the likelihood that EC HQs/EUDs use evaluation findings in its self-reflection, its interventions, and the dialogue within the European Commission, with other European Institutions, and with partner countries.

Each evaluation includes:

- A **preparatory phase**, during which the evaluation manager sets up the necessary institutional arrangements (the Inter Service Steering Group/Reference group, if relevant), defines the main features of the evaluation (roadmap for strategic evaluations, terms of reference), and selects the evaluators
- The **inception phase**, that aims at structuring the evaluation. The evaluation manager interacts with the evaluators in order to produce the evaluation design (reconstruction of the intervention logic and definition of evaluation questions and related judgement criteria and indicators, with identification of data collection tools and sources). The mapping and analysis of relevant spending (projects, programmes, budget support, etc.) and non-spending (policy dialogues, etc.) interventions, and the methodological proposal for the following phases, are part of this phase. The setting up of a consultation strategy is also required at this stage.
- **Implementation phase**. A phase during which the evaluators conduct the **research**. It can be broken down in two complementary phases:
- **Desk phase**: During this phase, desk work takes place in order to collect and analyse data, and coming up with preliminary answers to the evaluation questions and hypotheses that can guide the subsequent field work, whenever necessary.
- **Field phase**: field activities help in validating/rejecting preliminary answers to the evaluation questions and bring additional information and direct evidence.
- During this phase, the evaluation manager and the evaluation team interact in formal and informal meetings. When needed, the evaluation manager activates the Inter Service Steering Group/Reference group or other panels on the deliverables and to solve problems arising.
- **Synthesis and reporting phase.** This phase entails the analysis of the data collected during the desk and field phase (if organised) to finalise the answers to the evaluation questions, and prepare the synthesis report that includes the conclusions and recommendations of the evaluation. The Inter Service Steering Group/Reference group or other panels support the

[166] Refer to: https://myintracomm.ec.europa.eu/dg/devco/resources-procedures/it-information-systems/eval/Pages/index.aspx.

evaluation manager in assessing the quality of the draft deliverables in order to achieve their finalisation.

- A phase during which **dissemination** takes place. The final report, and additional deliverables if requested, is disseminated.
- The final phase during which the evaluation manager drafts the **action plan** (SWD, if relevant) and follows up.

**TABLE 16: THE MAIN PHASES IN DG NEAR EVALUATIONS WITH KEY MOMENTS AND DELIVERABLES, BY EVALUATION TYPE**

| Steps | Evaluation typologies | | | | |
|---|---|---|---|---|---|
| | Strategic | | | | Project/ Programme |
| | Major evaluations (Financial instrument; thematic impacting on policy revision) | Thematic | Country | Aid delivery | |
| **Strategic choice on what to evaluate and resources allocated** | | | | | |
| Evaluation plan | √ | √ | √ | √ | √ |
| Funding | √ | √ | √ | √ | √ |
| Appointment of an Evaluation manager | √ | √ | √ | √ | √ |
| **Preparatory** | | | | | |
| ISG set up | √ | √ | √ | √ | |
| Reference Group set up | | | | | √ |
| Roadmap prepared | √ | √ | √ | √ | |
| Roadmap published | √ | | | | |
| TOR prepared | √ | √ | √ | √ | √ |
| Procurement | √ | √ | √ | √ | √ |
| **Inception phase: structuring stage** | | | | | |
| Consultation strategy | √ | √ | √ | √ | √ |
| Kick-off meeting | √ | √ | √ | √ | √ |
| Inception report | √ | √ | √ | √ | √ |
| **Desk phase: data collection and analysis stage** | | | | | |
| Desk report, if foreseen | √ | √ | √ | √ | √ |
| **Field phase (if foreseen)** | | | | | |
| Field report, if foreseen | √ | √ | √ | √ | √ |
| **Synthesis phase** | | | | | |
| Public consultation | √ | | | | |
| Summary of public consultation | √ | | | | |
| Final report | √ | √ | √ | √ | √ |
| Publication | √ | √ | √ | √ | √ |
| **Dissemination and Follow up** | | | | | |
| Action plan | | √ | √ | √ | √ |
| Staff Working Document | √ | | | | |
| Follow-up on action plan | √ | √ | √ | √ | √ |

*Source: DG NEAR*

## 13.1. What to evaluate and resources allocated

### *13.1.1. Evaluation plans*

Following the requirement of the Commission Communication on strengthening the evaluation function[167], each evaluation function should elaborate a multi-annual evaluation plan. The Evaluation Plan is a part of the **Management Plan** of each DG and of each Delegation and thus announced publicly[168]. Procedures and responsibility for its approval are the same.

An Evaluation Plan indicates all **the evaluation activities** of an organisation: internal and external evaluations; ex ante, mid-term, and ex post evaluations; evaluations of all types of instruments and interventions, be they spending (e.g., projects, programmes) or non-spending (e.g., policy dialogue). It is drafted at various levels: the European Commission Evaluation Plan results from the plans of all DGs.

**DG NEAR Evaluation Plan is therefore a multi-annual plan,** including evaluation activities planned and implemented by Headquarters, following the principles of division of labour between the HQs, EUDs and the beneficiaries. It is updated on an annual basis and published as part of the DG NEAR Management Plan[169].

**BOX 43: PURPOSES OF EVALUATION PLANS FOR BOTH EC AND NATIONAL AUTHORITIES (WHEN RELEVANT)**

Evaluation plans help Headquarters, Delegations, and the beneficiaries adhere to **regulation mandates and guidelines requests** about evaluation and ensure that further **evaluation needs** of DG NEAR are fulfilled, by making evaluations results available when needed:

- to support preparation or modification of interventions and strategies
- to provide information for DG NEAR Annual Activity Report (AAR) and Joint DG DEVCO/NEAR Annual Report on financial cooperation
- to justify resource allocations
- to support negotiations and policy dialogue with partner countries
- to inform implementation choices.

Evaluation Plans also **inform stakeholders about what is being and will be evaluated**: European Institutions, Member States, the general public, other donors and partner countries' authorities, public and private organisations, the evaluation and development communities, and general population. DG NEAR strategic evaluations are listed in the European Commission Evaluation Plan.[170]

*Source: DG NEAR*

**BOX 44: EVALUATION PLAN PREPARATION: A COLLECTIVE ENDEAVOUR**

The evaluation plan has to be prepared in close consultation with the relevant stakeholders/users of the evaluation results.

In the case of the **DG NEAR HQs**, these are the country/regional programme units, the strategy and policy units, coordination unit, the audit, finance and contract units, the EU Delegations, etc.

In the case of the **EUD**s, these are the DG NEAR HQ's Monitoring & Evaluation service, the country teams/ regional programmes units, the programme managers in the EUDs and in consultation with the partner countries.

In the case of the **partner countries**, the relevant stakeholders are: EUDs, the line ministries, NIPAC officers, National coordinators for ENI, other donors active in the country, economic and social groups, representatives

[167] European Commission, 2013.

[168] In the future the evaluation plan should be integrated with the monitoring plan.

[169] More information on the DG NEAR plan is available at: http://ec.europa.eu/atwork/synthesis/amp/doc/near_mp_en.pdf.

[170] European Commission, 2015e.

of affected groups, etc.

*Source: DG NEAR*

Each **Delegation** and **operational unit** has to draft, report on and update their own annual Evaluation Plan. **DG NEAR HQ's Monitoring & Evaluation service** consolidates the different plans and supports their drafting and implementation.

**Beneficiaries**, especially under the indirect management mode, have to prepare an annual Evaluation Plan in consultation with the EUDs and publish it.

**BOX 45: STEPS IN DRAFTING AND MANAGING AN EVALUATION PLAN**

**FIRST STEP: IDENTIFY INSTITUTIONAL RESPONSIBILITY FOR DRAFTING AND MANAGING THE EVALUATION PLAN**

The first step is to identify a clear institutional responsibility for drafting and managing the Evaluation Plan and for keeping all relevant actors informed: in DG NEAR HQs operational Units, in Delegations and within the partner coordinating structure there must be an Evaluation Focal Point.

**SECOND STEP: IDENTIFY WHAT TO EVALUATE**

The Evaluation Focal Point can use the following **sources**:

- Regulations and other relevant instructions which envisage evaluations;
- DG NEAR policies, strategies, etc. linked to EU political priorities. Or directly linked to international commitments (i.e. SDGs, aid effectiveness, etc.)
- DG or Delegation activities. All European Commission actions must be periodically evaluated. EC HQs and Delegations need to schedule final evaluations for **projects/programmes above €5 million**. In addition, Delegations may ask for evaluations of particularly innovative actions, pilot actions, and sensitive actions (actions for which there have been implementation problems or conflicts or about which there has been controversy during action preparation) or themes on which there is little evidence.
- Requests from European Institutions, particularly the European Parliament, the Council, the Court of Auditors, etc.
- Top management requests for information may signal that there is a need for evidence on an intervention, group of interventions or theme.
- For Delegations, consultation with the national aid coordinator and with the partner country authorities.
- Interventions about which there are "new strategic decisions, implementation problems, indications coming from monitoring results, feedback on the REFIT programme, complaints from stakeholders, EU Pilot files, infringement procedures, audit reports" (European Commission, 2015c: 260).

These sources provide an indication of information needs. The Evaluation Plan Manager (that is the M&E focal point) should determine which ones can be satisfied by using information from the monitoring system, from ROMs, from statistical sources, or from existing studies and evaluations to satisfy these needs. These should not enter the Evaluation Plan.

The Evaluation Plan manager should **prioritise** the possible evaluations. Criteria are:

- **Interest** in the evidence from the DG top management or the Delegation management.
- **Need to use the evidence in mandated or important tasks**, such as the Mid-term Review, Annual Reports, other information material, preparation for next planning/programming period, development or amendment of a strategy, policy dialogue or negotiations with the partner country.
- **Feasibility within the timeframe and resources available.** It is important to also consider human resources: commitment and ownership of an evaluation are key for its utility. If it is impossible to fit the evaluation in staff's work programme, consider postponing or dropping the evaluation.
- Programme resilience vis-à-vis and management's willingness to accept **unexpected or unfavourable answers**—especially when evaluation reports have to be made public.
- **Interest** from national aid coordinator, partner country's authorities and feedback from

**stakeholders** (European Commission, 2015c: 260).

**THIRD STEP: CHOICE ON EVALUATION'S TYPE AND ON AVAILABLE RESOURCES**

For each evaluation, the Evaluation focal point needs to collect information in order to make choices:

- Whether the evaluation is going to be conducted internally (an option only possible when there are skilled and available human resources) or externally (which increases the amount of financial resources necessary).
- The amount of needed financial and human resources. These can be identified in full in the first stages of the evaluation, but the Evaluation Plan should at least indicate a range, based on experience.

**FORTH STEP: SCHEDULING THE EVALUATIONS**

For each evaluation, it is necessary to **identify the time when results are finally available** (including reviews and the preparation of the Staff Working Document if needed). From this, the Evaluation focal point can calculate the time when the preparation should start. The start date is calculated from the moment in which the preparation starts, and includes the time needed to prepare the evaluation and to select and contract consultants (when the evaluation is going to be conducted by external consultants).

For intervention evaluations the best timing is at the end or, better, **after the end of the intervention**, in order for achievements to materialise and be manifest.

Evaluation Plans should allow for time to support main reviews of the financial instruments or the preparation of the new financial framework and the new instruments, for which an impact assessment is also required. Delegations should also take into consideration additional **milestones** specific to the partner country (e.g., the national planning/programming cycle or the political cycle).

**FIFTH STEP: PUBLICATION**

Once the Plan is approved, the manager of each Evaluation Plan ensures its **publicity** vis-à-vis the relevant stakeholders (e.g. in the case of Headquarters, on the DG website, in addition to the publication of DG NEAR evaluations in the European Commission Evaluation Plan), and in ways that cater to other audiences, such as other donors and partner countries' authorities, public and private organisations, and the general population. Delegations publish their Evaluation Plan on their website, as well as in the consolidated DG NEAR evaluation plan.

**SIXTH STEP: MANAGING AN EVALUATION PLAN**

At this point, the difficult part begins. Managing an Evaluation Plan implies many tasks. First, the Evaluation focal point needs to ensure that the relevant Programme Manager **activates each evaluation** according to the schedule. Especially for multi-annual Evaluation Plans, or when many months have elapsed since final approval of the plan, or if there have been significant changes in the environment, this requires ensuring that the information need which gave rise to the evaluation is still present and that it still cannot be satisfied through existing data. Then, the manager needs to check whether the allotted resources are still available. An evaluation manager needs to be appointed and supported, especially in the first phases of the evaluation, when the evaluation needs are further developed into evaluation questions. After this, the evaluation manager drafts the evaluation roadmap (**BOX 48: THE EVALUATION ROADMAP**), if needed.

*Source: DG NEAR*

**BOX 46: CONTENT OF AN EVALUATION PLAN**

An evaluation plan must clearly state:

- the **responsibilities** for the plan: which entity (unit/Delegation) is responsible for drafting and updating it each year; for deciding in-year revisions of the plan, if necessary; for activating the units and individuals responsible for each evaluation; for ensuring that the resources are available and used;
- updated contact information for the **contact** point for the plan (European Commission, 2015c: 259)
- the **time frame** of the plan: the Better Regulation guidelines suggest a rolling plan covering at least 5 years for strategic evaluations. Activities should be well described for the first two years. For subsequent years, the plan is indicative, since the plan needs to adapt to emerging needs. For EU

Delegations and national authorities only an annual plan is required;

- an indication of the **resources** available for evaluations for each year. Resources include both **financial resources** and their **source** for managing the evaluations or conducting them (European Commission, 2015c: 257);[171]
- **complementarities and synergies** with the partner country' evaluation planning, if any exists (for Delegation evaluation plans) and with other donors' evaluation planning.
- key elements for **each evaluation**:[172]
- The **title**, **object** (a project/programme, a theme, a strategy, an aid modality), and indicative **scope** of the evaluation;
- the **purpose** of the evaluation;
- which **type of evaluation** it is (project, programme, thematic, etc.);
- whether it is going to be conducted **internall**y or by **external consultants;**
- the **resources** which will be available;
- **when** its results have to be available;
- the time when the evaluation process will **start;**
- **arrangements** and **responsibilities**, namely, which units will act as contracting authority and evaluation manager;
- whether it is a **joint evaluation** with the partner country or other donors.

*Source: DG NEAR*

The Evaluation Focal Point **updates the Plan each year**. The Evaluation Focal Point specifies the activities for the additional year(s), updates the information about on-going and completed evaluations, and checks whether there is the need to add new ones or to drop any evaluation which is not needed anymore. The manager also updates all the information regarding the Plan, starting from the available resources, taking into consideration the issues that arose during the implementation of the Evaluation Plan.

Throughout the year, finally, the Evaluation Plan manager should suggest in-year updating of the Plan if this is needed. In order to do this, the manager needs to promptly **identify new evaluation needs** that emerge from EU strategy alterations, changes in conditions within the partner countries, and changes in programmes. Further needs may arise from requests for information from Headquarters, management, stakeholders, and partner countries.

### *13.1.2. Appointment of an Evaluation manager*

**The quality and usefulness of an evaluation is as much the responsibility of the evaluator as it is of the evaluation manager**. The role of the evaluation manager is central to the evaluation.

Once the implementation of an evaluation is decided the first step is the appointment of the evaluation manager[173]. This person will normally be the relevant Programme manager responsible for the implementation/monitoring of the related Action.

In principle, the evaluation manager should not be (nor have been) involved in the planning/programming or implementation of the action(s) to be evaluated (in project, programme, budget support evaluations) or considered (in country/regional, sector, thematic, etc. evaluations). The unit where the evaluation manager

[171] While drafting the evaluation plan, the required human resources should be duly considered.

[172] For evaluations still in the pipeline, this information is only indicative. The Evaluation Plan manager gradually updates it as they it is specified.

[173] For a detailed list of its tasks, please refer to **TABLE 4: EVALUATION ACTORS AND RELEVANT KEY RESPONSIBILITIES.**

works should not formally depend from the Unit responsible for the action. This is meant to protect the **credibility** and **impartiality** of the evaluation.

In reality, both in action evaluations at HQs and EUDs, the evaluation manager is or has been involved in the action and/or operates in the unit/section responsible for it. **This does not necessarily detract from the utility and quality of the evaluation**.

- First, there is great value in this arrangement, because it **increases ownership**, the amount of implicit and explicit **knowledge** evaluators tap into, and the probability of **use** of evaluation results. This coincides with the primary purpose of evaluation at the European Commission: supporting Commission services in improving their work.
- Second, even in these conditions, there are many ways in which the **credibility** and **impartiality** of the evaluation can be protected, including a **closer intervention of a group supporting the quality of the evaluation** (a Reference Group or, when it is foreseen, an Inter Service Steering Group), the creation of an **expert panel** to review deliverables, and **stakeholders' involvement in the evaluation** (for example, via presentations to stakeholder groups). It is also good to remember that action evaluations may always be reviewed when Headquarters launches a **strategic evaluation or a systematic review**: in these cases, problems with credibility and impartiality become evident. Thus, a frequent practice of using action evaluations in reviews and in strategic evaluations provides a strong incentive to guarantee credibility and impartiality.

Evaluation managers need knowledge of DG NEAR operations and of key evaluation concepts. This knowledge enables them to follow the evaluation from the preparation phase to follow-up on the action plan. Particularly, a knowledge of key concepts in evaluation methods helps in determining the resources needed for an evaluation, selecting the evaluation team, interacting with evaluators during the choice and combination of evaluation methods during the inception phase, pursuing a dialogue with the evaluators, and interpreting the results on the basis of the limits and potentialities of the main approaches and methods.

The Delegations and the national authorities (whenever relevant) evaluation managers can count on the support of **DG NEAR HQ's Monitoring & Evaluation Service**.

The interaction between the contracting authority, represented by the evaluation manager, and the evaluators is a delicate one.

#### *13.1.3. Funding of the evaluation*

DG NEAR recommends that an appropriate budget should be earmarked in the planning/programming stage (within the Action Document, or the Action Programme, or in the framework of the European Integration Facility (EIF) for Enlargement countries) for carrying out evaluations. These evaluations would complement the recurrent monitoring missions (which can only ascertain whether we are doing things right), and assess whether or not we achieved results.

Concerning the funding of the evaluations, the following principles have to be respected:

- For all planned evaluations, irrespective of the management mode, an envelope approximately equivalent to 2% of the total value of the annual and/or multi-annual programmes should be set aside. The amount should reflect the actual evaluations to be implemented, considering that for certain types of actions (like budget support operations, where the evaluation would be under the HQ responsibility, the EIF itself, and small actions, where no specific evaluations should be contemplated and not implemented through direct management) evaluations will be financed from DG NEAR HQ's Evaluation budget.
- For programmes implemented through indirect management, evaluations would be contracted by the national authorities (unless the Financing Agreement formally foresees the opposite). Should

the quality/independence of the evaluations be questionable, an evaluation could be launched by the EC using Technical and Administrative Assistance Facility (ATA) funds (under the EUD leadership) or the central budget.

- For Cross-border Cooperation (CBC) programmes:
    - Implemented through direct management: whereas a 2% quota cannot be kept aside since the programmes covering activities until 2017 are already adopted, the evaluations will be covered by a Technical Assistance, which is already foreseen.
    - Implemented through indirect management: it is proposed that an evaluation covering all programmes of the 7 year period be carried out either at mid-term or at the end of the 7 years and financed form DG NEAR HQ's Evaluation budget.
    - Implemented through shared management with Member States: the responsibility to carry out evaluations lays with the Managing Authority of the programme and shall be financed from programme technical assistance budget; however, the European Commission can at any moment, launch evaluation or monitoring of the programme or of a part thereof.
- Strategic evaluations of budget support, in general carried out addressing several interventions in a given country, will be normally covered by DG NEAR HQ's Monitoring & Evaluation service. Individual evaluations of budget support interventions may be carried out at Delegation level using specific templates of terms of reference drafted by DEVCO budget support unit.

Determining the cost of an evaluation is a delicate task. Usually, it is suggested using a criterion of proportionality to the intervention. In reality, a better guide is offered by **the amount and type of work needed to answer pressing questions**. The reason why it is suggested to have few evaluation questions is precisely to make the best use of scarce human and financial resources. Evaluations of relatively small intervention which are innovative, very controversial, or which are key for future strategies may require more resources than evaluations of large interventions that can mostly rely on desk research. The evaluation manager needs to form an idea of the costs of both desk and field activities. For example, fieldwork, even when performed by local staff, tends to be expensive, especially if it is to be performed in many countries. Some techniques for data analysis are very labour intensive.

The financial provisions included in the Evaluation roadmap (and ToR) need to be based on the **elaboration of a task-budget**[174], in which the evaluation manager sets a first provision of working days for each team member allocated by evaluation phase and type of task. The task-budget also includes the average of daily subsistence allowances and travel costs for each evaluation phase. In addition to this task-budget, a calculation needs to be done, if relevant, for: translation, printing, logistics (for the organisation of seminars or other types of venues), and publication and dissemination activities.

The main variables are: the geographical diversity needed for fieldwork, labour intensiveness of needed research activities, especially if original, the quality of existing data and availability of data from monitoring systems, the size of target population (and of the control groups if applicable), the size and diversity of the evaluation team, and the products which are asked (translations, for example, may be expensive).[175]

[174] Annex 17 provides a standard model of an evaluation's task-budget.

[175] "The cost of an evaluation can vary a great deal depending on a number of variables, such as: size and duration of the intervention, scope and complexity of the intervention, geographical diversity, size and nature of the stakeholders/target population/beneficiaries, quality of monitoring systems in place, data readily available and methods foreseen. Generally, the more is asked from the contractor the higher the costs" (European Commission, 2015c: 257).

## 13.2. Preparatory phase

### *13.2.1. Setting up an Inter-service steering group/Reference group*

The request for setting up an ISG/RG is made by the evaluation manager in charge of the evaluation through a note.

The **Inter-service steering group (ISG)** is requested by the Better Regulation to be set up for each **evaluation managed by HQs**. It should be composed of a minimum of three members. In the case of strategic evaluations launched by DG NEAR, there should be at least one representative from the unit leading the evaluation (DG NEAR HQ's Monitoring & Evaluation service, in order to provide relevant methodological knowledge and to protect the autonomy and credibility of the evaluation, and/or any operational units in DG NEAR).

**BOX 47: PURPOSES OF AN INTER-SERVICE STEERING GROUP AS PER THE BETTER REGULATION, AND BY EXTENSION, PURPOSES OF A REFERENCE GROUP**

The Inter-Service Steering Group/Reference group:

- **Steer the evaluation exercise in all key phases** (roadmap (only in the case of ISG), Terms of reference, selection of the evaluation team, consultation, inception, desk, field and synthesis and reporting phases, Staff Working Document (only in the case of ISG). The evaluation manager[176] steers the ISG and the Reference group and is supported in its function by ISG and RG members.
- **Provide input and information** to the evaluation team and **demonstrate an open and transparent approach to critically analysing performance and delivery**. Indeed, the different knowledge and perspectives emerging from their members must ensure a sound data collection process and data analyses. The mix of policy/operational experience and evaluation expertise should prevent particular biases from taking over the evaluation and encourage critical judgements to be made.
- In the case of ISG, they mobilise the institutional, thematic, and methodological knowledge available in the various DGs of the Commission that are interested in an evaluation.
- **Ensure the respect of principles** set up in **BOX 3: EVALUATION PRINCIPLES**. In this framework: the **independence of evaluators** and the **impartiality and usefulness**[177] of both the evaluation process and the final deliverable(s), are particularly important.
- **Provide quality control** on the different deliverables. The evaluation manager, as lead of the ISG and RG, consolidates the comments to be sent to the evaluation team.
- Endorse the different deliverables.
- Elaborate the quality assessment grid[178]
- **Ensure a proper follow-up action plan** after the completion of the evaluation

*Source: DG NEAR, adapted from the Better Regulation*

In the case of **project/programme evaluations**, a **reference group (RG)** will have to be set up.

When setting up the ISG or the RG, the possibility of inviting representatives from the main stakeholders can be considered. For evaluation in DG NEAR, this might mean representatives of other donors and of

[176] In the case of project/programme evaluations, the evaluation manager usually corresponds to the programme manager.

[177] "It used to be common to regard the use of evaluation as being confined to acting on recommendations and final reports. It is now understood that evaluation use can be supported and occurs throughout an evaluation. Process use should involve stakeholders in evaluation thinking from the beginning. There are evaluations where the conclusions and recommendations are rejected but stakeholders, especially those involved in the steering committee find the evaluation useful. It can help them to clarify their own thinking and understanding and spark off innovative ideas for improvements. Promoting dialogue during the course of an evaluation is likely to ensure that when stakeholders receive reports they will be better prepared and receptive" (European Commission, 2013: 59).

[178] A template of such grid is enclosed in Annex 20.

partner countries' authorities and civil society. The evaluation manager has the freedom to decide and suggest the stakeholders and to invite them to nominate their representatives. However, it should be noted that the group should not be too big in order to be operational and manageable (max 15 people).

### *13.2.2. Evaluation roadmap*[179]

According to the Better Regulation Package, **the roadmap should be prepared for all major evaluations** (these are those referring to the financial instrument or those having impact on a policy revision), as well as on thematic, sector, aid modality evaluations launched by the DG NEAR HQ's Monitoring & Evaluation service. **Project/programme evaluations managed by Delegations,** or thematic evaluations launched at local level, **do not need one**. Evaluation roadmaps are to be published on the central website of the Commission only for evaluations carried out by DG NEAR HQ's Monitoring & Evaluation service.

The evaluation manager, together with the ISG is responsible for drafting the evaluation roadmap. The **preparation of the roadmap is an essential process** in the evaluation design as it is the first stage of definition of the subject, purpose and scope of the evaluation, the issues to be addressed and the evidence that will be gathered. In addition, it is aimed at informing wider public about the launch of a forthcoming evaluation. It should be, therefore;

- Prepared well in advance (e.g. once the evaluation plan for the current year is approved);
- Published on European Commission Better Regulation page for major evaluations and DG NEAR evaluation page)[180], thus providing access for the public. The aim is to inform stakeholders and citizens about new initiatives as well as about evaluations. **Publication of the roadmap provides greater transparency** and enables stakeholders to provide their feedback.

**BOX 48: THE EVALUATION ROADMAP**

**For each strategic evaluation**, an evaluation roadmap, compliant with the template provided by the Secretariat General must be published centrally on Europa website (.....) and on DG NEAR website.

On published roadmaps, stakeholders are invited to give feedback that could feed into the further preparatory process, which usually starts 4 weeks following publication (So if stakeholders want their input to be considered, they should ideally provide feedback within this 4 week period).

The **roadmap will indicate**:

- The subject of the evaluation, its purpose and scope (including draft evaluation questions and where relevant other tasks);
- The background and original objectives of the intervention being evaluated (including a short description of how these were expected to be achieved);
- The questions covering the five mandatory evaluation criteria (or an explanation about why not all five criteria are addressed);
- The evidence base for the evaluation, covering both data already available and data which will be collected during the evaluation, including key elements of the consultation strategy;
- Key expected dates including, the start of the evaluation, timing of consultation and publication of the final report;
- Contact details allowing stakeholders to provide feedback.

[179] Only applicable to evaluations launched by the DG NEAR HQ's Monitoring & Evaluation service. A template and example of an evaluation roadmap can be seen at: https://encrypted.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0ahUKEwi7xuad6ZXOAhWDXBQKHSHTBfgQFgglMAE&url=http%3A%2F%2Fec.europa.eu%2Fsmart-regulation%2Froadmaps%2Fdocs%2F2017_near_001_evaluation_ipa2_en.pdf&usg=AFQjCNHIHhOenZS6I_nDrzrbZgtY_LUUpQ.

[180] The roadmaps of the major evaluations (according to the definition of the Better Regulation Package) are published on the European Commission Better Regulation page (http://ec.europa.eu/smart-regulation/roadmaps/index_en.htm). The roadmaps of other evaluations are published on each internet page of the Contracting Authority that commissioned the evaluation (DG NEAR (http://ec.europa.eu/enlargement/news_corner/key-documents/index_en.htm?key_document=08012624887bedda) or EUDs or national authorities.

The final version of the evaluation roadmap will be signed off by the Director General and published on the central website. Where relevant, it may also be published on the related policy pages of the author service.

*Source: Evaluation Commission, 2015c: 264*

There is no deadline for feedback from any interested party/stakeholder on the roadmap, though after four weeks from its publication the terms of reference can be drafted. However, it is advisable that the evaluation manager collects possible feedback well in advance before starting preparing the Terms of Reference.

**BOX 49: TIPS FOR THE PREPARATION OF THE EVALUATION ROADMAP**

During the preparation of the evaluation plan, evaluation managers should:

- **Reconstruct the intervention logic** of the subject of the evaluation as a first step to define sound Evaluation questions. This work, reconstruction of the IL and elaboration of the EQs, would be the backbone of the roadmap and further elaboration of ToR.
- Make sure to **include questions they really need answers** to, along with questions relevant to mandatory evaluation criteria (refer to **BOX 4: MANDATORY EVALUATION CRITERIA AS PER THE BETTER REGULATION GUIDELINES, WITH TYPICAL EXAMPLES OF QUESTIONS**)
- **Select carefully evaluation questions**, motivating the choice, as the Better Regulation guidelines suggest, making them relevant for the policy objectives pursued by the EU, as reflected in DG NEAR management plan.
- Involve senior managers.[181]

*Source: DG NEAR*

### *13.2.3. Terms of reference*

The Terms of Reference set out what is expected by the contractor in the case of evaluations contracted out to external consultants.

The following requirements are stated in the Better Regulation:

- Terms of Reference must be established for all evaluations outsourced to contractors.
- The Quality Assessment criteria need to be included as an annex to all Terms of Reference.
- The ISG/RG has to be consulted on a draft of the Terms of Reference before it is published.
- All external studies must respect the Commission Visual Identity and be registered in the Inter-institutional database of studies from the planning stage and then, if not confidential, published in EU Bookshop. This requirement doesn't apply to operational evaluations launched by Delegations.

Terms of Reference for evaluation face a specific challenge. They have to be precise enough to **guide** both evaluators and the evaluation manager. At the same time, they have to be flexible enough to leave space for **creativity and autonomy**. Creativity and autonomy are important to ensure the credibility and quality of evaluation. This suggests building the Terms of Reference in terms of the **questions** that have to be addressed, while avoiding referring to the specific methods to address them, except in particular cases[182]. A good practice is to ask bidders to propose **combinations of methods and techniques**, thus increasing the robustness of results.

[181] Better Regulation guidelines recommend this as an instrument to increase instrumental use ("Senior management should be actively involved to ensure the use evaluation results in the decision making process." European Commission, 2015c: 220. Similar suggestions in European Commission, 2013: 57).

[182] For example, evaluations inquiring about attribution might require an experimental approach or evaluations on sensitive issues or on interventions which engendered controversies or conflicts in the partner country might require a participatory approach

As in the case of the intervention logic diagram's elaboration**, it is also recommended to ask for the advice of a member of DG NEAR HQ's Monitoring & Evaluation service once a first ToR's draft is available**.

**BOX 50: CONTENT OF THE TERMS OF REFERENCE** [183]

In the Terms of Reference, the evaluation manager specifies:

- The **title** of the evaluation.
- Background/context analysis
- The **type of evaluation and its main purposes, objectives and (temporal, thematic, geographical, etc.) scope**. The scope of the evaluation: the interventions that will be the object of the evaluation, or, at least criteria to determine this key element at a later stage (for example, in thematic and strategic evaluations) and the relevant time-span.
- Reconstruction of the intervention logic of the subject of the evaluation
- The **evaluation questions** the evaluation will need to address. Their relevant judgement criteria and indicators, and identification of data collection tools and sources, are expected to be finalised by the evaluation team and agreed upon by the Evaluation manager.
- A list of relevant publicly **available documentation** (complete with functional links).
- The **timeline** of the evaluation, focusing on the time when the results are expected and the final deliverables are to be delivered. This requires careful consideration of data availability and of the methods compatible with the evaluation questions.[184]
- **Information on the management responsibility** (Unit/section in charge of providing the Evaluation manager) and on the **members of the ISG/RG.**
- Minimum requirements in terms of **experience** and **composition of the evaluation team**. This is particularly important when requesting a combination of different methods and approaches.[185] The evaluation manager must strike a balance between ensuring that the evaluation team possesses the relevant skills and avoiding a restriction of the market or barriers to entry. This might be a concern especially when requiring local experts in smaller countries or in countries with little evaluation tradition or requiring experts with a sufficient knowledge of the local context and/or of the language. It is good practice to require that teams include thematic expertise, research experience, research methods, and evaluation expertise. It is also good practice to allow that teams are diverse by field or geographical region in which individual team members have acquired their experience.
- In principle, evaluators should not be or have been involved in the intervention and they should be independent from the unit that manages or has managed the intervention.[186]
- Financial resources.
- **Intended uses** and **users** of evaluation results.
- The **deliverables** expected for dissemination: in addition to the basic evaluation reports and executive summaries, it is good practice to include when relevant non-technical reports, materials for dissemination (translations of syntheses and reports in the partner country language, visual

[183] The EVAL Module (https://myintracomm.ec.europa.eu/dg/devco/resources-procedures/it-information-systems/eval/Pages/index.aspx) will be gradually fed by Programme managers with documents pertaining to evaluation exercises. In this regard, it will be possible to read and get inspiration from different sets of documents, including ToR.

[184] The timeline is extremely important when it has been decided to use an experimental method: in this case the evaluation must start at the same time that the project is designed.

[185] "One of the biggest problems that those who manage or commission evaluation face is how to put together a suitable team or mix of competencies that may properly come from all these traditions" (European Commission, 2013: 15). Especially when asking for a combination of methods, this issue is relevant and the resources must accommodate for larger and more experienced teams which are able and willing to work across approaches and evaluation traditions.

[186] For an exploration of the concept of independence, see Picciotto, 2013.

| documentation, videos), and presentations to the main stakeholders and a wider public. |
|---|

*Source: DG NEAR*

#### 13.2.3.1. Evaluation questions

**Evaluation questions** determine what the evaluation focuses on, the approaches and methods it uses, and the findings it produces.

Evaluation questions:

- Should not exceed a manageable number, which can be answered within the limits of available resources and time. Ideally, not more than ten.
- Are linked to evaluation criteria. Refer to **BOX 4: MANDATORY EVALUATION CRITERIA AS PER THE BETTER REGULATION GUIDELINES, WITH TYPICAL EXAMPLES OF QUESTIONS** and **BOX 5: OTHER EVALUATION CRITERIA, WITH EXAMPLES OF TYPICAL QUESTIONS.**
- Request knowledge in order to **use** it[187] Indeed, they are asked by stakeholders which have some degree of responsibility for or an interest in a given policy and/or action. As such, they should be defined in close consultation with the key stakeholders of the evaluation.
- Include the request for an informed, evidence-based, value **judgment**.
- **Emerge from the reconstructed intervention logic.** While acknowledging that each evaluation exercise has its own scope (an evaluation can be focused on a given level of the intervention logic, on certain parts of different levels, etc.), which is linked also to the type of evaluation, evaluation questions should usually:
- Include elements that are within the sphere of control (from inputs to outputs[188]) of the EC/EUD, national authorities
- Include elements (e.g., results and societal needs), which are at least partially out of the control of the organisation responsible for the intervention to be evaluated, be it EC/EUD, national authorities. These pertain mostly to outcomes, which are directly influenced by an EC/EUD/national authority intervention[189]. Depending on the type of evaluation, these can also refer to intermediate and long-term impacts to allow an assessment of the structural changes to which the intervention has possibly indirectly influenced (contributed to).
- Focus on the main elements of the thematic/country/regional intervention (to allow an appropriate assessment of the implementation and the actual outputs and outcomes, and the efficient use of the inputs provided
- Consider issues of the management structure, the type of instruments used and the way they have been used
- Present the strategy envisaged for additional data collection and data analysis.

**Identifying evaluation questions is a shared responsibility**. During the planning phases of evaluation, **contracting authorities, programme managers** and **evaluation managers** identify and express the

[187] The first use that comes to mind is the "instrumental" use: whenever the results of an evaluation are used to modify an intervention. This is a comparatively rare form of use. There are other uses as well (some commendable, others less): to learn about the intervention; to create a level field among partners; to further a position in negotiations; to justify decisions which have already been made; to feed the democratic debate about a policy field; to demonstrate results to the general public, to democratic institutions, or to funders of interventions; or to fulfil norms and other obligations.

[188] For a definition of the levels of the results chain, refer to **FIGURE 4 : A SIMPLIFIED (LINEAR) DEPICTION OF THE RESULT'S CHAIN OF AN ACTION**.

[189] Ibidem.

questions keeping into consideration requests from DG NEAR's top management, the European Parliament or the Council, and involving intended **users of evaluation** and **stakeholders** (e.g., partner country authorities and stakeholders within partner countries[190]).

**Internal and external evaluators** finalise the evaluation questions during the **inception phase[191], by adding judgement criteria and indicators,** and build the **evaluation methodology** (data collection tools and analysis) on them, in interaction with evaluation managers.

**BOX 51: ELEMENTS OF AN EVALUATION QUESTION**

An evaluation question is composed of different elements. They are to be defined in a sequenced manner.

**The evaluation question headline.**

This is the element that first focuses the question's scope.

For example:

- To what extent has EU used its available instruments in a way that enhances complementarity in support of the overall EU goals of a healthy environment, sound natural resource management and strong environmental and climate governance in developing countries?

**Rationale**

The rationale provides the justification for the evaluation question's choice and linked to this it presents its main scope.

For example (linked to the above example):

*The question evaluates the extent to which there has been a synergy and complementarity between environment support funded under geographic instruments (usually where environment/climate change is a focal sector) and the ENRTP[192].*

*The question covers environment and climate change actions financed by both geographic instruments and the ENRTP in countries with environment/climate change as a focal or non-focal sector for the EU (i.e. in the 11 case countries selected for this evaluation).*

*The rationale for selecting this area for an evaluation question is that:*

- *The original intention of the ENRTP was to deal with important environment and climate change issues that could not be covered (as well) by geographic instruments – often these are global issues or have a global (or transboundary) aspect;*
- *There are opportunities for synergies between the two instrument types, which should have led to additional benefits;*
- *There is also a risk that actions under ENRTP merely substituted or duplicated efforts under geographic instruments;*

*The first judgement criterion assesses the extent to which actions funded under ENRTP enabled EU to address environment or climate issues, which it would be difficult to address sufficiently under the geographical instruments. The second judgement criterion assesses the extent to which actions financed by ENRTP and by geographic instruments have benefitted from/been strengthened by each other. The last judgement criterion is similar to the second criterion, but focuses on the complementarity between ENRTP and actions funded by member states and other donors.*

**Evaluation criteria covered**

**Evaluation criteria connected to the evaluation questions provide a guide to identify judgment criteria, but do not coincide with them** (see JCs examples below).

[190] Refer to **13.3.1.1 CONSULTATION STRATEGY** on the importance of stakeholders' involvement.

[191] The evaluation exercise may provide elements which prompt evaluators, evaluation managers, or intended users of evaluation to revise evaluation questions or add new ones, within the limits of available resources (time, staff, and financial resources), during the desk phase.

[192] Former EC Thematic Programme. Environment and Sustainable. Management of Natural. Resources including Energy.

For example (linked to the above example), the evaluation question covered:

*Relevance – looks at the extent which environment support has been successful in responding to the continuum of national, regional and global priorities*

*Efficiency – looks at the extent to which the variety of instruments have achieved a synergy that is cost effective*

*Coherence – looks at the extent to which the thematic and geographic instruments have been coherent amongst themselves and also with other broader EU policies*

*Added value – looks at the extent to which the combination of instruments has delivered greater value together than they could have done alone*

### Judgment criteria

**Judgment is the essential feature of evaluation**: after assessing the achievements of an intervention and understanding how they are produced, evaluation asks how good they are and whether they are good enough. **Judgments are the sole responsibility of the evaluators.** Judgments are based on **values**[193].

Evaluators develop **judgment criteria**, which specify an aspect of the merits or success of the evaluated intervention. It is used to answer an evaluation question positively or negatively[194]. To some extent, they can be defined as well as sub-questions of the main headline. Their number (per evaluation question) should be manageable.

**Evaluation criteria connected to the evaluation questions provide a guide to identify judgment criteria, but do not coincide with them**.

For example, if we refer to the evaluation headline referred to above, it has two parts:

- What is being judged: "EU support".
- The way of judging: Has it "… enhanced complementarity…'[195]

The judgement criteria are meant to develop and specify the second part of the question. For this question, the Judgment criteria are:

- JC 1. ENRTP has enabled the EU to address environment and climate change issues, which could/would not have been better, or equally well, addressed through its geographical instruments
- JC 2. Environment and climate change actions financed by geographic instruments have benefitted from the ENRTP (e.g. from research, environmental reforms or the climate policy changes triggered by ENRTP)
- JC 3. ENRTP actions have benefitted from complementary action financed through geographic instruments (through either project or budget support modalities) and through non-EU cooperation

### Indicators.

Judgment criteria are operationalized through **indicators**. As such, indicators can only be defined once the judgement criteria are drafted. In certain cases it could be advisable to ask for the evaluators to first develop the judgement criteria, and once these are agreed by the ISG/RG, ask them to define the indicators and identify their sources of information.

Indicators specify in advance which data are to be collected. They thus help to focus the data collection process[196].

Here it is important to highlight that **indicators defined during planning/programming need to be included to the extent possible in evaluations (into specific evaluation questions).** As mentioned

[193] Evaluation theories differ on the basis of the way they assign value to an intervention: whether achievements correspond to objectives, whether achievements reach external standards, or whether achievements fit in with stakeholders' notion of success (European Commission, 2013). Different theories originate different approaches and methods.

[194] Methodological basis for evaluation, Joint Evaluation Unit, *Op. Cit.*.

[195] Adapted from Methodological basis for evaluation, Joint Evaluation Unit, *Op. Cit.*.

[196] Ibidem.

elsewhere, depending on the scope of the evaluation exercise, additional indicators can be identified, but priority should be given to existing planning/programming indicators. Also, **indicators can emerge from relevant monitoring databases and performance frameworks** (the EU results framework in the case of ENI and the IPA II performance framework). Indeed, this would help the **evaluation feeding into corporate reporting.**

For example (linked to the above example):

- For JC 1, indicators are:
- Level of difference between ENRTP actions financed via geographical instruments in relation to their focus, approach, scope and implementation
- Degree of ENRTP's use in actions that could not have been done equally well or better using available geographic instruments.
- For JC 2, indicators are:
- Degree of environment and climate change actions' (implemented under geographic instruments) use of information and research provided by ENRTP actions in their country/region
- Degree of sustainability of environment and climate change actions implemented under geographic instruments due to ENRTP policy related interventions in their country/region
- For JC 3, indicators are:
- Degree of ENRTP actions' consideration of actions financed by EU under its geographical instruments
- Degree of ENRTP actions' consideration of actions financed by other EU Member States and/or other donors

**Sources of information.**

Each indicator can consider **one or more sources of information**.

For example, the sources for indicator '*Actions under ENRTP differ in their focus, approach, scope and implementation from actions that can be, and are, implemented under geographical instruments' are:*

Documents/data:

- Data from the ENRTP and geographic instruments actions' inventory (list of EU spending actions)
- Country strategy papers for the planning/programming period 2007-2013
- Progress reports, completion reports, evaluations, ROM, EAMR, annual activity reports for sample actions
- Surveys

Interviews with:

- EC HQs, EUDs
- National partners
- Action's staff
- other

*Source: DG NEAR*

**BOX 52: TIPS AND CRITERIA FOR SELECTING EVALUATION QUESTIONS**

The interaction between key stakeholders may produce quite an impressive list of questions that would need an answer. Some of these, however, such as those relating to inputs, costs, and outputs may find an answer from **monitoring** or other **administrative data**. If these are available and are considered to be robust, they can be to the extent possible disregarded as such.

Some evaluation questions may find good answers in **systematic reviews** of literature and evaluations. These are evaluation questions about interventions in sectors that have already been the object of intense evaluation (e.g., education or social policy). Reviews are a comparatively fast and frugal way of obtaining solid and useful answers. They require very skilled teams, combining methodological expertise in performing reviews and in-depth knowledge of the subject matter and the policy field.

When choosing evaluation questions, the following criteria should be taken into consideration:

- Genuine **willingness to know the answer and to use the knowledge**. An organisation only truly sustains the effort (in terms of time and human, financial, and organisational resources) if it believes that it will use the results. Otherwise, there might be a compliance attitude or, worse, the risk of leaning on the evaluator to provide the desired answers. This suggests using great care in mandating evaluation questions that should guide all evaluations. When confronted with evaluation questions they do not share, programme managers and evaluation contracting authorities might oppose resistance. This should be taken into consideration especially when trying to build in-country evaluation capacity.
- **Feasibility**: whether the answer to the question may be found in a reasonable amount of time and within the limits of available resources.
- **Resources** (time, financial resources, composition of the evaluation team, time and quality of the individuals who manage the evaluation, available data) **must be appropriate to the tasks**.
- Action management's willingness and strength to **accept unexpected or unfavourable answers**—especially when evaluation reports have to be made public.
- **Ownership**: the organisation responsible for the evaluation (which acts as the contracting authority or which ensures management of the evaluation) must **cooperate in formulating the evaluation question** and in requesting stakeholders' inputs.
- **Consensus**: evaluation managers and evaluators will give higher priority to evaluation questions that are relevant to the higher number of stakeholders.
- Choose **as few questions as possible**. Overburdening an evaluation with questions and criteria results in poor evaluation quality.

*Source: DG NEAR*

**BOX 53: TYPE OF EVALUATION QUESTIONS, WITH TYPICAL EXAMPLES**

There are various **types of evaluation questions**: [197]

**Descriptive questions**

They **ask what has happened** and require evaluators to define, observe, and measure change, often from the point of view of various stakeholders. These questions pertain to positive and negative changes, be they expected or unexpected, directly or indirectly linked to the intervention. **They are always to be asked**. In fact, **they create the basis for all other analyses**.[198]

Typical examples are

[197] Evaluation questions differ substantially from the questions that drive research activities, research questions, which may have any scope. Research questions are not necessarily aimed at facilitating decision-making. They do not explicitly provide judgments and are usually formulated by the researcher, rather than by organisations.

[198] In the OECD-DAC framework, these questions pertain to one of the criteria, Impact. The Better Regulation framework states that impact must always be assessed: "Evaluations and Fitness Checks should also always assess the economic, social and environmental impacts of EU interventions" (European Commission, 2015c: 71).

- *Which are the changes that have occurred?*
- *Which positive and negative changes are being produced?*
- *When the intervention works at its best, what does it produce?*
- *How do changes differ for each area/sector/affected group?*

**Causal questions**

They are the **essence of evaluation**: they connect changes to the intervention. Causal questions shed light on whether an intervention works and on how it works, for whom, and under which circumstances. Research aiming at answering them shows the path through which EU interventions interact with public and private actors in partner countries to produce outputs, outcomes, and impacts. Therefore, they produce knowledge that can be used to improve interventions, to identify indicators, to understand problems and fix them, and to launch new, effective initiatives in the future.

There are different types of causal questions**:**

First, there are those which require **explanations**:

- *How has the intervention elicited changes? Through which mechanisms? Who has been affected?*
- *Under which circumstances does the intervention work?*

Second, causal questions inquire about **attribution** and **how much of the observed changes** can be attributed to the intervention:

- *Has the intervention caused the observed changes?*
- *How much of the observed changes is the consequence of the intervention (rather than of intervening factors)? For example, how many new enterprises have support to SME created?*

Third, they relate to **contribution:**

- *Which factors interacted to produce the observed changes?*
- *Which role has the intervention played in determining the changes?*

Generally speaking, DG NEAR evaluations do concentrate on the first and third types. The contribution approach is also the philosophy behind the EU Results Framework and the IPA Performance framework.

**Normative questions**

They **ask how an intervention fares against a criterion.** Regulations request criteria—but dot not rule out the possibility of adding other criteria (refer to **FIGURE 2: THE INTERVENTION LOGIC AND THE KEY 7 (5 BEING MANDATORY) EVALUATION CRITERIA**). Guidelines, such as Better Regulation (European Commission, 2015b) and the OECD DAC principles also lay down sets of criteria and request evaluation managers and intended users to select the criteria that are most relevant for the task at hand. For examples of questions for each of these criteria, please refer to **BOX 4: MANDATORY EVALUATION CRITERIA AS PER THE BETTER REGULATION GUIDELINES, WITH TYPICAL EXAMPLES OF QUESTIONS** and **BOX 5: OTHER EVALUATION CRITERIA, WITH EXAMPLES OF TYPICAL QUESTIONS.**

When selecting criteria, evaluation managers always have to **explicitly state the rationale for their choice** in the documents accompanying the start of an evaluation. Later, they should also **request that evaluation reports include this rationale.** In addition, evaluation designs can devote a reasonable amount of resources to questions that need fewer efforts to be answered in that particular evaluation.[199]

*Source: DG NEAR*

[199] "The degree of analysis conducted for each criterion will depend on the intervention being evaluated, the timing of the evaluation and the reliability of the data (proportionality). Often this will mean that for some criteria new data will need to be collected, analysed and triangulated with other findings; whilst for others, a short summary can be presented based on existing reports and information", European Commission, 2015c: 71.

## Possible questions for each evaluation type

**There is no primer for evaluation questions**. They depend on the **scope** and **level** of the evaluation (an intervention evaluation differs from the evaluation of a country strategy), on the **type of intervention** (sector reforms, infrastructural programmes, or interventions **aiming** at individuals, e.g., the unemployed prompt different questions and different methods), on the **purpose** of the evaluation, on **evaluative knowledge** already possessed by decision-makers, and on the **knowledge needs** of intended users.

Evaluation managers and evaluators must use knowledge about the interventions, the sector, and the country(ies), and must investigate the evaluation needs of the Commission, the partner country, and the main stakeholders. **BOX 54: TYPE OF EVALUATION QUESTIONS, BY MAIN TYPE OF EVALUATION LAUNCHED BY DG NEAR** shows which type of questions it is possible to ask for each type of evaluation[200].

**BOX 54: TYPE OF EVALUATION QUESTIONS, BY MAIN TYPE OF EVALUATION LAUNCHED BY DG NEAR**

**Project evaluation**.

**Most evaluation questions (and evaluation methods to answer them) are fit for project evaluation**: *Descriptive questions* about positive and negative changes, *causal questions* about *attribution* and *explanation*, and *normative questions* about *efficiency* can find better answers at this level than at more strategic levels, especially for projects that are focused on individual activities.

Other questions, such as normative questions about EU-added value or impact, on the contrary, may encounter conceptual difficulties due to the scale of the project. In addition, since impact is defined as the *intermediate and long-term* consequence of an intervention, impact questions require that enough time has elapsed.

In specifying questions for project evaluations, evaluation managers must **consider the scale, complexity, and nature of the project**. For example, *attribution questions* can be answered by counterfactual methods when the nature of the project allows for building counterfactuals—that is, if the intervention is implemented homogeneously and if it is feasible to identify comparable individuals or organisations or territorial units which are eligible for the intervention but do not participate in it and collect data about them[201].

Questions about *impact* on socio-economic variables such as economic growth can be asked only for projects accounting for a large percentage of a country public intervention in a sector.

The same applies to *environmental sustainability*: questions about changes in a river water quality may be asked only for projects which affect waste water treatment for a substantial proportion of the population of the river basin, the use of water for most of agricultural production in that river basin, or industrial and urban waste management in that same territory. For smaller project in sectors such as water and sanitation, housing, industrial and agricultural development, it is recommended to assess how the facilities created or restructured use water resources and how the changes in industrial or agricultural activity increase (or decrease) water consumption and increase (or decrease) pollution at the project level.

Questions about project-related *coordination* are relevant only for projects that have many different components concurring to the same objective.

**Budget Support**.[202]

The **specific features of budget support influence the type of questions** that can be asked. First, direct achievements refer to institution building: the consequences of the agreements in terms of changes in policy design, influence on sector reforms, and the introduction of new procedures. Changes induced in citizens' lives and in the business environment, instead, can be conceptualised in terms of *contribution.* In addition, the most relevant types of questions seem to be descriptive questions and normative questions

[200] Annex 19 associates each group of questions with possible methods and approaches to answer them.

[201] As mentioned earlier, attribution analysis is very rare in DG NEAR's evaluation portfolio. The focus is on contribution analysis.

[202] As mentioned earlier, budget support has its own evaluation methodological approach. The evaluation methodology can be found at: http://www.oecd.org/dac/evaluation/dcdndep/Methodological%20approach%20BS%20evaluations%20Sept%202012%20_with%20cover%20Thi.pdf.

about effectiveness, impact, relevance, and institutional sustainability.

**Country strategy evaluations**.

**Most questions are appropriate for country strategy evaluations**. They can be asked both as refers as the strategy as a whole and referring to particular sectors or instruments of interest: *descriptive questions*, causal questions about contribution, and normative question relating to relevance, effectiveness, impact, utility, internal and external coherence, EU-value added, complementarity, coordination, equity and acceptability. *Causal questions* about attribution and explanation and normative questions about efficiency and sustainability are more appropriate for individual strategy components and can form the basis of comparisons.

**Thematic evaluations.**

All types of questions can be asked for thematic evaluations. These evaluations may answer questions about aggregate achievements (for example, which are the achievements of transportation interventions in IPA II countries over a given period of time?). **Some questions (e.g., *descriptive questions, causal questions about contribution, and normative questions* about relevance, effectiveness, impact, sustainability, equity, and acceptability) are best suited to this task.**

**The greatest value of thematic evaluations, however, lies in answering questions implying comparisons among interventions of various types.** For example:

- *Under which circumstances have implementers kept costs down to the lowest level without compromising achievements?*
- *Which instruments work better in spurring competitiveness in SMEs and why?*
- *Under which circumstance has the EU value-added been highest?*
- *Which type of interventions best addressed the needs of Roma people and how?*
- *Which implementation arrangements allowed to best redressing gender imbalances?*
- *Under which circumstances do interventions engender the highest support for public sector reforms?*
- *Which types of interventions prove to produce the most sustainable achievements from the institutional and financial points of view?*

Systematic reviews and realist syntheses of existing interventions' evaluations and studies are particularly appropriate to thematic evaluations and complement original evaluative research.

**Other strategic evaluations**.

**All types of questions are appropriate for strategic evaluations**. *Descriptive questions and normative questions* about effectiveness, impact, sustainability, utility, and equity may be answered both at the level of the general strategy and by comparing strategy tools or the way the instruments have been used in each country or for various contexts. *Causal questions and normative questions* about efficiency provide best results when asked at this second level, by comparing different tools or different areas.

**Within the framework of DG NEAR interventions**, aimed at IPA II countries and at sustaining the Neighbourhood policy, **the greatest value for strategic evaluations comes from normative questions about EU-added value**. For example:

- *What is the added value of EU interventions, compared to what could be achieved by the partner countries?*
- *Which areas do not require the involvement of EU support because they are well covered by other donors?*
- *What is the added value of the EU financing compared to funds provided by IFIs and/or national financial institutions?*[203]

Also:

- *Internal coherence:*

[203] Evaluation roadmap for DG NEAR strategic evaluation on SME competitiveness: http://ec.europa.eu/enlargement/pdf/financial_assistance/phare/evaluation/2015/20151028-evaluation_roadmap_competitiveness.pdf.

- *How do the various components of the strategy combine?*
- *How do direct support interventions for SME combine with sector reforms? How do interventions under a given strategy interact in producing/redressing gender inequality?*

▪ *Coordination:*
- *Which organisational arrangements ensure coordination among the various components of the strategy?*
- *How is institutional continuity ensured during the time span and the various phases of the strategy?*

▪ *External coherence:*
- *To what extent is the EU assistance coherent with interventions by other international actors?*
- *To what extent is EU assistance coherent with other EU interventions in related fields?*

▪ *Complementarity*
- *to which extent partner policies and actions were complementary to the EU-supported intervention?*

▪ *Acceptability*
- *To what extent has the EU assistance enhanced the coherence and visibility of EU aid, and promoted innovative approaches?* [204]

*Source: DG NEAR*

**TABLE 17: TYPE OF EVALUATION QUESTIONS, BY MAIN TYPE OF EVALUATION LAUNCHED BY DG NEAR**

| Type of evaluation question | | Type of evaluation | | | | |
|---|---|---|---|---|---|---|
| | | Strategic | | | | Action (Project/Programme) |
| | | Thematic/ sector | Instrument | Aid modality (i.e. Budget Support) | Country | |
| **Descriptive** | | √ | √ | √ | √ | √ |
| **Causal** | Explanation | √ | √ | √ | √ | √ |
| | Attribution | | | | | √ |
| | Contribution | √ | √ | √ | √ | √ |
| **Normative** | | | | | | |
| *Mandatory criteria as per the Better Regulation* | Relevance | √ | √ | √ | √ | √ |
| | Effectiveness | √ | √ | √ | √ | √ |
| | Efficiency | √ | √ | | | √ |
| | Coherence | √ | √ | √ | √ | √ |
| | EU-added value | √ | √ | √ | √ | √ |
| Other criteria | Impact | √ | √ | √ | √ | √ |
| | Sustainability | √ | √ | √ | √ | √ |
| | Utility | √ | √ | √ | √ | √ |
| | Complementarity | √ | √ | √ | √ | √ |
| | Coordination | √ | √ | √ | √ | √ |
| | Equity | √ | √ | √ | √ | √ |
| | Acceptability | √ | √ | | √ | √ |

*Source: DG NEAR*

[204] Ibidem.

### 13.3. Implementation

Even though strategic and project/programme evaluations differ in resources and timing, as well as in methodology205, the present chapter provides a sound picture of the main phases an evaluation can have. Depending on the specificities of the evaluations launched, these can be adapted.

#### *13.3.1. Inception phase*

This phase aims at:

- Clarifying the issues of the evaluation. Indeed, the **inception phase** starts with the **kick-off meeting**. The meeting has the purpose to "arrive at a clear shared understanding of what is required by the contracting authority".
- Revise and/or reconstruct the intervention logic of the object of the evaluation
- Based on the latter, finalise the evaluation questions and present the overall framework of the analysis.
- Agree on the work plan, budget and evaluation team.

As such, **it represents the backbone of the evaluation**, and outlines the set of objectives against which the relevant intervention(s) will be assessed.

**BOX 55: INCEPTION PHASE ACTIVITIES, WITH EVALUATOR AND EVALUATION MANAGERS' RESPECTIVE RESPONSIBILITIES**

The inception phase includes all the activities leading to **the finalisation of the evaluation design**:

**Assessment of data consistency and quality**.

- The **evaluation manager** supports the evaluator in identifying datasets, in accessing, and in familiarising with them.
- The **evaluator** collects all available data and assesses its quality **relative to the evaluation questions**. Although the evaluation might ask an assessment of the quality of monitoring data, this activity focuses only on whether the data is available and good enough to answer the questions.
- Collection of action/country/region/sector/other documents.
- The **evaluation manager** provides all the documentation relative to the interventions, including documents that are not publicly available and confidential materials, whenever appropriate.
- The **evaluators** seek documents that are not available to the Commission.

Drafting of the **evaluation design**:

- The evaluators
  - Complete a preliminary **review of literature** on: EU/EC relevant policy and normative framework, EU/partner country/region relations, mapping of relevant spending and non-spending interventions, etc.
  - Based on this review, they finalise the reconstruction of the intervention logic
  - Propose the evaluation design (based on the elements already indicated in the ToR):
  - **Finalise the evaluation questions**, including related rationale, judgment criteria, indicators and sources[206], as well as a clear explanation of the rationale of their choice (especially when

[205] Strategic evaluations launched by DG NEAR HQ's Monitoring & Evaluation services, launched via the framework contract COM 2015, need to comply with the Evaluation Methodology for European Commission External Assistance, developed by DEVCO.

[206] Depending on the evaluation scope:

- Qualitative data is collected mostly through:

evaluation questions mandated by the Better Regulation guidelines or by the OECD DAC criteria have been excluded).

This activity might need additional meetings with the relevant units and DGs at Headquarters and in the country(ies), at the Delegation, with the national aid coordinator and with the partner country authorities and representatives of civil society. The scope of this activity depends on the scope of the evaluation

- The **evaluation manager** updates the information that was in the Terms of Reference and assists the evaluator in accessing key informants and data.
- The **evaluator** finalises the questions (defining the judgement criteria, indicators, sources of information and data collection tools), in interaction with the evaluation manager.
- If appropriate, first meetings in partner countries can be organised in order to revise the evaluation questions.
- Propose the **approaches** to be taken, the **methods and techniques** to collect and analyse data and to arrive at judgments, for the desk, field and synthesis phases.
- If relevant, and this is often the case for strategic evaluations, propose (if they were not already specified in the terms of reference or in the evaluators' proposal) or finalise a **sample[207] of case studies** to be assessed in order to answer the evaluation questions.
- For the choice of case studies, it is suggested that **evaluation manager** ensures the preliminary identification of a minimum or maximum number of case studies and identification of general criteria to select them (refer to **BOX 56: EXAMPLES OF CASE STUDIES SELECTION CRITERIA**).
- **Evaluation managers** and **evaluators**, with the support of the ISG/RG, agree on the choice (and on the standardized structure for data collection, analysis and reporting).
- In cases in which an evaluation does not have a proper desk phase (finalised with a desk report), the inception report must also detail the focus of the fieldwork by proposing , in the case of strategic evaluations (thematic, regional, aid modality) a limited number of countries, along with their selection criteria (i.e. geographical coverage, income development level, share of EU financial commitment compared to total sectorial/regional commitment, sector specific considerations (their weight in the intervention logic), representativeness of beneficiaries, type of aid modality,

---

  - EU policy and strategy documents- such as Council Conclusions, Communications, Agreements with partner countries and regions, etc.
  - Country/region official documents, such as National Development strategies
- EU country/region/sector intervention documents:
  - Action documents, etc.
  - Previous evaluation work, especially ex ante evaluations. During intervention development, options are discarded and choices are made—sometimes risky choices. Ex ante evaluations should keep track of both discarded options and bold choices: these are good candidates for proposing evaluations and for finding out which the doubts and intentions were during intervention development. Interventions and instruments over which there has been controversy during intervention's development or in negotiations also provide good, pressing evaluation questions.
  - Monitoring findings. Evidence from monitoring activities prompt evaluation questions.
- Interviews, meetings and focus groups with key stakeholders and informants, direct observation, surveys. Stakeholders and key informants: each actor in Neighbourhood and Accession policy interventions has its peculiar standpoint and, therefore, its own needs for knowledge about the results of interventions—knowledge they can use for improving the interventions as evidence to support their positions in negotiations, to learn from past experience, and for demonstrating results to the European Parliament, to the countries, to Member States, and to the general public.
- Beneficiaries' own reports, websites, etc.
- Other donors' documents
- Independent studies, evaluations
- Quantitative data mainly consist of:
- Statistical data on country/sector profile from WB, IMF, OECD, Central Banks, etc.
- Country's aid data from national database/OECD/UNDP
- EC projects – namely, financial data by sector, country and year, and project outputs (if quantifiable)
- Other main donors' actions (budget commitments).

[207] It is possible to include in the sample cases that have not participated in work relative to that theme, in order to provide for comparisons.

availability of information, etc.)

- At this stage it might also not be possible to identify all the case studies: in this case, **the evaluators** must provide a justified list of the criteria they will use and of the time frame within which the choice will be finalised.
- The evaluation manager:
  - Discusses the evaluation design and agrees on it, if relevant involving the Inter Service Steering Group/Reference group, the expert panel, or the stakeholder group. Involving stakeholders in the evaluation design helps achieve credibility of the evaluation.
  - Should hold meetings on the evaluation questions and on the methods to answer them, focusing on the limits of and requirements for each method and on how the methods interact. Evaluation managers should request original field and desk research and ensure that evaluators seek and use additional data sources throughout research activities. In this phase, evaluation managers should request that for each indicator of an evaluation question, specific data sources are defined.
  - Ensures that the evaluators explain all points in a clear and simple way and that, whenever possible, they accept suggestions from stakeholders.

Proposals for **dissemination products** in addition to the final evaluation report.

Based on the ToR requirements, the **evaluator** and the **evaluation manager** agree upon these.

**Preparation of all the following steps**, on the basis of the evaluation design. The **evaluators** and the **evaluation manager** cooperate on this.

The **evaluation manager** provides information on the intervention(s) and contacts with relevant actors in the partner country(ies).

*Source: DG NEAR*

As shown in **BOX 55: INCEPTION PHASE ACTIVITIES, WITH EVALUATOR AND EVALUATION MANAGERS' RESPECTIVE RESPONSIBILITIES, evaluators develop the evaluation design in close cooperation with the evaluation manager**. This does not imply that evaluation managers should become experts in research methods or data analysis, but requires that they possess a knowledge of some basic features of the main evaluation approaches, namely which questions an approach can answer to, which resources it needs, which are its requirements in terms of timing and expertise.[208] & [209]

**No approach and no method are appropriate for all interventions and for all evaluation questions**. Often, only a combination of approaches and a design based on mixed methods can provide robust results. This is all the more true for the wide scope of interventions and sectors covered by DG NEAR in both ENI and IPA II countries. The relevant criteria in choosing approaches and methods are:

- Ensure the approach is **appropriate** to the evaluation question, to the nature of the intervention, and to the type of evaluation.[210] Some types of evaluations, namely country strategy evaluations, thematic evaluations, and other strategic evaluations have complex evaluation objects and, therefore, **always need combinations of approaches**.

[208] For example, an example is that of participatory approaches. These require skills and resources—for example for reaching the stakeholders, for performing the necessary fieldwork.

[209] Very synthetic descriptions of some of the most common evaluation approaches which are appropriate to evaluating DG NEAR interventions are provided in Annex 18. Descriptions draw upon Marchesi, G., Tagle, L. and Befani, B., 2011.

It should also be kept in mind, however, that the field is in constant movement, and that methods and approaches are continuously being proposed and refined. Evaluation uses research to produce policy-relevant knowledge. As such, it uses all possible methods in economics, in social and political sciences, in humanities, in environmental and gender studies. It thrives on creativity and dissent. New evaluation approaches, methods, and techniques are constantly suggested, tried, and used. Existing ones are used in new sectors, to answer different questions, and to deliver valuable knowledge in fields for which they were hitherto considered unfit. Any attempt to a final categorisation is vain and rife with controversy.

[210] Annex 19 provides abstract examples of correspondence between evaluation questions and type of evaluation.

- Use the **simplest and most frugal** methods and techniques that allow reaching quality results. Fieldwork, of course, can be replaced if feasible by online surveys, phone interviews, etc.
- Minimise the time and effort requested from respondents and data providers in the partner country(ies).
- Consider **feasibility**: the timeframe, data availability and resources.

The evaluation manager can find support from:

- DG NEAR HQ's Monitoring & Evaluation service
- Methodological guidance available online and in publications[211].
- Searches of evaluative literature on the internet or in journals (Evaluation, New Directions for Evaluation).
- The expert panel or the Inter Service Steering Group/Reference group
- Discussion groups (evaltalk, the website of the outcome mapping community)
- Evaluation training.

**BOX 56: EXAMPLES OF CASE STUDIES SELECTION CRITERIA**

Case studies can be considered as in-depth, self-contained studies embedded within a larger study or evaluation. Case studies are usually used to generate and analyse data of a retrospective nature about a particular entity (the case-study object), for example, an action[212] or a set of actions, an aid modality, a sector, a country, an institution, etc. [213]

Case studies aim at providing a view of the EU results in the different components of a particular (.e. sectorial/country/regional) strategy and they are to be chosen in such a way to ensure at best EU's support representativeness.

Case studies cannot cover the entire scope of an evaluation and as such they are to be considered as one of the elements of data collection and analysis, to be complemented with information collected from other sources.

Criteria commonly used, to be adapted on a case-by-case basis, are:

[211] Excellent methodological guidance, geared towards a wider audience than evaluators, is provided by these, very diverse, sources:

- Better Evaluation (http://betterevaluation.org) provides an online resource on all aspects of evaluation. Its numerous entries provide synthetic explanations, bibliographies, and rapid access to online resources.
- EvalPartners (http://www.evalpartners.org/) provides toolkits and a virtual library, among other useful resources.
- European Commission (2013), EVALSED: The resource for the Evaluation of Socio-Economic Development. The guidance combines texts, sourcebooks and a glossary. It is conceived as a guide for evaluation of socio-economic interventions. It, therefore, focuses on evaluation of spending actions rather than on analysis of policy mixes, such as the ones which form the focus of DG NEAR activity, whereby spending actions coexist with institution creation and building and reform of legal systems.
- My M&E (http://mymande.org) is an online platform providing various types of resources (videos, webinars, training) on country-led evaluation, equity based evaluation, and development evaluation. It is the hub for a diverse and vibrant community.
- Stern, E. (2015) Impact Evaluation. A Guide for Commissioners and Managers. London: bond. This publication synthesizes and makes easily accessible for non-evaluation specialists approaches and methods for evaluating impacts and draws upon the report Broadening the Range of Designs and Methods for Impact Evaluations.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., Befani, B. (2012), Broadening the Range of Designs and Methods for Impact Evaluations. DFID, Working Paper 38. This publication explores various approaches to evaluate impacts. Although it is part of a highly learned and specialized evaluation debate, it provides useful information that evaluation managers may use.
- Other resources, focused on specific sectors, are under construction, some funded by the European Union. One example is Impact Europe (http://impacteurope.eu), which aims at developing a toolkit about evaluation of counter-violent radicalization interventions.

[212] A case study is not meant to be though an intervention's evaluation.

[213] Adapted from Evaluating EU activities. a practical guide for the commission services, DB Budget Evaluation Unit, July 2004.

- **Representativeness of the different strategic components (i.e. sectorial clusters for example) identified in the intervention logic**. It could be useful to identify at least one case study per area of intervention.
- **Availability of documentation**. The documentation made available for a given action should include: preparatory and identification studies, action documents, action reports (inception, progress reports, final reports), monitoring reports, evaluation reports, minutes of action's steering group meetings and other internal documents. In the case of a sectorial case study for example, relevant actions would also need to have the above mentioned documents, and these would need to be complemented by EU policy and normative documents pertaining to a given sector, and for a given country/region, as well as partner countries sectorial policy and strategy documents.
- **The geographical coverage** for strategic evaluations (mainly regional and thematic).

  In thematic evaluations, a number of considerations may guide the choice of countries:
  - A certain balance of sub-regions
  - A certain balance based on the particular weight of the importance and/or concentration of EU support
  - A certain balance based on the type of EU/partner country strategic framework
  - Practical considerations in respect to spatial concentration and accessibility within one mission, to optimize the efficiency of travel and information gathering.
- **The size of the actions**. And their share compared to that of EU sectorial/country/regional total financial commitment.
- **The duration of actions**: those spanning over 2 or 3 years should be preferred.
- **Elimination of recently started actions**, unless they present innovative characteristics.
- **Mix of funding and implementation instruments**.

Furthermore, a number of secondary criteria may be defined with the aim of achieving a certain balance between:

- Type of support (Services, works, grants)
- Country specific interventions, regional and global interventions,
- Implementing partners (Government Organizations, Public Agencies or institutions, Multilateral and Regional Organizations, Bilateral Donors, Private Sector, EU and partner country(ies) NGOs, Community-based Organizations, Consulting Firms, Universities, Research Institutions, etc.)
- Beneficiaries (Government Organizations, Public Agencies or institutions, Multilateral and Regional Organizations, Bilateral Donors, Private Sector, EU and partner country(ies) NGOs, Community-based Organizations, Consulting Firms, Universities, Research Institutions, etc.)

*Source: DG NEAR*

#### 13.3.1.1. Consultation strategy

Understanding stakeholder expectations, and their priorities and information needs, is crucial in quality evaluations. The **evaluations must be informed and informing exercises**. Key stakeholders need therefore to be consulted before the launch of an evaluation exercise and during the different phases of such an exercise.

**Stakeholders** are all those actors that at any level are interested in the action or policy: from European institutions to partner countries authorities, to socio-economic groups in partner countries, particularly those who represent the intended or potential beneficiaries of the intervention. They should be involved in evaluation, to a proportionate degree and with appropriate modalities. They can be involved in various

ways: they can express evaluation needs, provide information[214], or contribute to ensure the credibility of an evaluation by participating in Inter-service Steering Groups/Reference groups or other forms.

There are various forms of **stakeholders' involvement in the evaluation**, apart from their participation in ISG/RG as mentioned under 13.2. Existing or ad hoc **stakeholders' group** can follow the entire evaluation process or be involved in key moments, such as the presentation of the method or the presentation of preliminary findings. These groups increase the credibility of an evaluation especially in conflictive situations or on controversial issues.

**BOX 57: TYPE OF STAKEHOLDERS TO BE INVOLVED, WHEN RELEVANT, IN EVALUATIONS[215]**

Among the key stakeholders we usually find **partner country authorities**, whose information needs and points of view should form part of evaluations. Under indirect management, they have the responsibility to launch and make possible evaluations. In all cases, they should provide data and information, facilitate evaluators' access to data and informants, and use evaluation findings in decision-making and negotiations.

**Potential and actual beneficiaries** of interventions are to be involved as providers of data. They should also have access to evaluation findings, through appropriate arrangements. Whenever it is possible, it is recommended to use participatory methods in order to involve beneficiaries (e.g., individuals who are affected by the interventions) in the evaluation exercise so that they can provide data and judgements in relation to specific issues.

**Expert panels** comprise thematic or evaluation experts who are external to the organisations involved. They usually review deliverables. The purposes are to protect the autonomy and credibility of the evaluation by guaranteeing a high profile methodological interaction with the evaluator. They are most useful when innovating on methods and approaches. It is good practice to allow the evaluators to indicate the experts or to jointly identify them.

An **advisory group** is an expert panel proposed and managed by the evaluation team. It usually discusses methodological issues with the evaluation team and reviews deliverables before they are sent to the contracting authority.

*Source: DG NEAR*

Internal evaluators and evaluation managers, though, may also use other opportunities, benefitting from the fact that in these cases they do not need to ask directly about evaluation questions, but can infer what the needs are from the discussions:

- During **action steering committees**. This allows linking the choice with the current debate on implementation. It reaches, however, only the stakeholders who sit in the committees.
- During **meetings in which the results of other evaluations are presented**, relative to that country, territory, or sector. This allows getting in touch with organisations and actors who are usually not part of the committees, including scholars, journalists, civil society.
- Finally, **formal consultations[216]** might be a source, for very large evaluations that grant the effort. It is better, however, to use the formal consultation at a later stage to validate results, rather than collecting evaluation questions.

[214] Evaluators have an ethical obligation to maintain the identity of respondents and of surveyed individuals confidential—even from the contracting authority. The evaluator manager should request that the evaluators divulge data in an aggregated and anonymous way, that they keep the data in a safe repository with all due cautions (e.g., coding interviews or individuals in databases and keeping the codes in a separate and safe place), and that they obtain explicit and written consensus from individuals before recording or taking pictures.

[215] Expert panels and advisory groups are not foreseen though under the global ToR of the framework contract COM 2015 (to be used in all strategic evaluations (those launched by DG NEAR HQ's Monitoring & Evaluation Service), and can therefore not be organised for evaluations contracted through this channel.

[216] See European Commission, 2015b.

**For major evaluations**, according to the Better Regulation, in the inception phase of the evaluation **the evaluation manager should prepare a consultation strategy to be discussed and agreed by the ISSG**, complying with the European Commission requirements for public consultation.

A simple consultation strategy should **identify and target relevant stakeholders** and evidence must be developed for each evaluation. Key elements of the consultation strategy should be **outlined in the evaluation roadmap**. The consultation activity must fulfil the Commission's minimum standards for public consultation, as outlines in the Better Regulation package (see Box below).

**BOX 58: GENERAL PRINCIPLES AND MINIMUM STANDARDS FOR CONSULTATION**

**Relations with stakeholders are governed by four general principles**:

1) Participation: Adopt an inclusive approach by consulting as widely as possible;
2) Openness and Accountability: Make the consultation process and how it has affected policy making transparent to those involved and to the general public;
3) Effectiveness: Consult at a time where stakeholder views can still make a difference, respect proportionality and specific restraints;
4) Coherence: Ensure consistency of consultation processes across all services as well as evaluation, review and quality control.

These principles are **complemented by five Minimum Standards** that all consultations have to respect:

A. Clear content of the consultation process ('Clarity'): All communication and the consultation document itself should be clear, concise and include all necessary information to facilitate responses;
B. Consultation of target groups ('Targeting'): When defining the target group(s) in a consultation process, the Commission should ensure that all relevant parties have an opportunity to express their opinions;
C. Publication: The Commission should ensure adequate awareness-raising publicity and adapt its communication channels to meet the needs of all target audiences. Without excluding other communication tools, (open public) consultations should be published on the internet and announced at the "single access point";
D. Time limits for participation ('Consultation period'): The Commission should provide sufficient time for planning and responses to invitations and written contributions;
E. Acknowledgement of feedback ('Feedback'): Receipt of contributions should be acknowledged and contributions published. Publication of contributions on the "single access point" replaces a separate acknowledgment if published within 15 working days. Results of (open public) consultations should be published and displayed on websites linked to the "single access point" on the Internet and adequate feedback given on how the results of the consultation have been taken into account.

*Source: DG NEAR*

Stakeholders must be able to provide feedback on each roadmap and draft evaluation report[217] (for major evaluations, according to the Better Regulation), or any legislative and policy proposals that will be adopted by the EC College and/or any draft implementing and delegated acts.

[217] It should be published on the DG NEAR/EUD/national authorities' web site, ensuring a 12-week internet-based public consultation.

**BOX 59: TYPES OF STAKEHOLDERS FEEDBACK**

**Feedback** in evaluation presents peculiar issues: it must increase quality while not infringing on the evaluator's autonomy. There are three types of feedback:

- Feedback on **techniques**, **methods**, **approaches, and ways to conduct the research**. These comments should primarily refer to what has been agreed in the evaluation design and subsequent changes. Evaluators accept these comments or provide an explanation of why they do not.
- Feedback on the **substance** of the document, the **findings**, and the **conclusions** should be expressed in a way that is respectful of the autonomy of the evaluator. If absolutely necessary, the evaluation manager can further ask for clarifications on specific points and, at that point, evaluators need to respond.
- Feedback concerning the **clarity** of writing or the accessibility to a non-technical public needs to be taken into consideration by the evaluator.

*Source: DG NEAR*

The involvement of key stakeholders since the early phases is also **key for their actual uptake once the evaluation exercise is finalised**: the elaboration of a management response and the establishment of effective follow-up processes on the recommendations are crucial in this regard.

At the end of the process, a report outlining the overall results of the consultation work and providing feedback (synopsis report) must be published on the consultation website and, where applicable, added as an annex to the evaluation report. Such a report would also provide an occasion to summarise relevant feedbacks received in parallel.

This phase ends with the elaboration and approval of an Inception Report.

**BOX 60: CONTENT OF THE INCEPTION REPORT**

The **Inception Report** includes:

- Introduction
  - Objectives and scope of the evaluation
  - Structure of the report
- A synthesis of all activities conducted
- Analysis of the subject of the evaluation's framework
  - Reconstruction of the *intended* intervention logic
- The proposed EQs (EQ's heading, judgement criteria and indicators, and relevant sources of information and data collection tools)
- Methodology for the remaining phases
  - Key methodological elements for each of the phases
    - It can include: The proposed consultation strategy
  - Overall approach for the desk phase
  - A proposal of a sample of countries/interventions subject to in depth desk review and field work, if any
  - Limitations
- Work plan
- The conclusions of the kick-off meeting
- A draft outline of the final synthesis report amending if necessary that included in the ToR
- A list of proposed dissemination materials

The Inception Report is accompanied by a statement of the Inter-Service Steering Group/reference group

highlighting any relevant issue

*Source: DG NEAR*

### *13.3.2. Desk phase*

This phase aims at:

- Completing the data collection and analysis and identifying information gaps
- Providing a preliminary answer to the EQs. Based on these, identify preliminary hypotheses and assumptions to be tested in the field
- Refine if needed the data analysis methods and detail potential limitations of the analyses
- Discussing potential amendments to the selection of interventions and/or case studies identified during the inception phase
- Proposing the methodology to carry out the field visits.

The **evaluative desk review phase** may last many months, depending on the scope and complexity of the evaluation.

**BOX 61: DESK PHASE ACTIVITIES, WITH EVALUATORS AND EVALUATION MANAGERS' RESPECTIVE RESPONSIBILITIES**

The evaluators will:

- **Complete the analysis of the intervention logic**. A first analysis of the intervention logic, allowing for its revision/reconstruction, was already done in the inception phase. For the completion of the desk report, a more complete analysis will be elaborated. In this framework, an in-depth documentary analysis, directly linked to the Evaluation Questions and relevant Judgement Criteria and indicators, will be done.
- For the selected interventions and/or case studies, the evaluators will proceed to the **analysis of relevant data.** This will also be directly following the information needs to answering to the Evaluation questions.
- **Undertake interviews, mainly at EC HQs' level.** Interviews will aim at gathering information and understanding on both i) sector/country/region policy and strategy definition and ii) implementation issues. Interviews, as other sources of information, will be delimited to the evaluation questions.

  The list of key stakeholders will be made taking into account several criteria, such as:

- **Prioritisation** in relation to:
  - Whether the person/institution is crucial for the selected intervention/sector because directly involved either in the implementation or as partner
  - Whether the person/institution is either a key informant for the intervention/sector without having been directly involved
  - Whether the person/institution is a key informant and could provide an useful insight at sector level
- The inclusion of **different type of stakeholders**, in particular:
  - EC and EUD staff (in HQs and in the field phase)
  - National and local governments (in HQs and in the field phase)
  - Main representatives of civil society (in HQs and in the field phase)
  - Other donors (in HQs and in the field phase)
- Based on the latters, evaluators will **draft a preliminary answer to the EQs** and identify a preliminary set of assumptions by EQ to be tested in the field.
- Define a **detailed field phase methodological approach**. The proposal will also detail the list of

stakeholders to be consulted during the field phase and planned visit timetables.

- Draft a desk report.
- Present preliminary findings at EC's HQs

The evaluation manager will:

- Support the evaluators in **completing data collection.** Access to data includes the usability of data and some assistance in understanding how the data is collected and by whom and how to interpret it. This holds true especially for monitoring and financial data.
- **Provide contacts** with relevant stakeholders in the partner country(ies)
- ***Discuss* the desk report**, particularly the draft answers to the evaluation questions and the proposed field approach involving the ISG/RG, the expert panel, and/or the stakeholder group.
- In addition to the draft desk report presentation meeting, the evaluation manager should **hold formal or informal meetings** to discuss particular issues, if needed

*Source: DG NEAR*

This phase requires that the evaluation manager uses care, attention, and ingenuity. All aspects of the evaluation must be creatively managed: issues may refer, in exceptional cases, to the emerging of new evaluation questions, to the need to alter evaluation questions as a consequence of changing strategic priorities with the European Commission or in response to intractable problems, and to strategies for data collection and analysis.

In this phase, the evaluation manager's need for guidance on evaluation approaches and methods intensifies. The strategies to cope with this are the same than during the inception phase.

This phase ends with the elaboration and approval of a Desk Report.

**BOX 62: CONTENT OF THE DESK REPORT**

The **Desk Report** includes:

- Introduction
- Background and key methodological elements
- Preliminary answers to the evaluation questions
- Approach and Methodology for the next phases
  - Overall approach for the field phase, including hypotheses to be tested
  - Remaining work for the synthesis phase
- Work plan

*Source: DG NEAR*

### *13.3.1. Field phase*

The field phase aims at:

- Completing the data collection. The fieldwork is meant to complete what is found in the inception and desk reviews. Data collected has to allow providing information to each evaluation questions' relevant indicator, so as to provide an answer to the judgment criteria and to the Evaluation Question. Therefore, this generally requires that field visits are standardized through the use of the same structure of questionnaires and same type of questions, fully in line with the Evaluation questions.
- Linked to the previous bullet, contributing to answer to the evaluation questions
- Validating or revising the preliminary hypotheses and assumptions formulated in the desk report
- Assessing whether there is need for further research and interviews to prepare the synthesis report, and in particular the conclusions and recommendation chapter.

The field visits are generally organized around two types of data to be collected:

- Data concerning overall EU country/region/sector policy and strategy
- Data concerning implementation issues (could be case studies on specific actions, sectorial analysis also through implemented actions, etc.)

The weight between the two levels of data collection depends on the final selection of the evaluation questions and judgment criteria, which will make clear the type of data still to be collected. Data collection focuses on the issues necessary for answering to the evaluation questions.

**BOX 63: FIELD PHASE ACTIVITIES, WITH EVALUATORS AND EVALUATION MANAGERS' RESPECTIVE RESPONSIBILITIES**

The evaluators will:

- **Undertake interviews in the field.** The evaluation questions and approach proposed for the field visits require the possibility of collecting data from different stakeholders and key informants. The same philosophy as for the desk phase interviews applies (refer to **BOX 61: DESK PHASE ACTIVITIES, WITH EVALUATORS AND EVALUATION MANAGERS' RESPECTIVE RESPONSIBILITIES**).
- The ***time available*** for the visits, the ***logistics*** of the visits and the ***people availability***
- A first **list of stakeholders and key informants** should be based on the desk analysis carried out and should be elaborated together with the evaluation manager and EUD staff. The list should be then finalized following the final planning for the field visits.
- **Draft a field report/note**. When possible, in order not to dedicate too many human and financial resources to intermediate reporting, evaluators can be asked to elaborate a **power point presentation** instead of an actual field report/note.
- **Present preliminary findings** in the field and at EC's HQs.

The evaluation manager will:

- Provides an outline for the field report/note/PowerPoint
- Intervene in facilitating fieldwork and access to data.
- The evaluation manager **discusses the** field report/note/PowerPoint involving the ISG/RG, the expert panel, and/or the stakeholder group.

*Source: DG NEAR*

### *13.3.1. Synthesis phase*

This phase entails the analysis of the data collected during both the desk and the field visits to complete the answers to the evaluation questions, and the preparation of the synthesis report that includes the final conclusions and recommendations of the study.

#### 13.3.1.1. Final evaluation report

In the final phase, evaluators synthesise their findings, judgments, and conclusions in a final evaluation report (see **BOX 64: CONTENT OF THE FINAL REPORT**).

**BOX 64: CONTENT OF THE FINAL REPORT**[218]

The **Final Report** includes:

- Executive summary
- Introduction
- Key methodological steps, including:
  - Description of all evaluation activities, focusing on the choice of evaluation questions, approaches and methods, limitations to the research (including those linked to data quality and if relevant, to monitoring systems), problems encountered and solutions implemented, and explanation of any difference between the evaluation design and the final report

[218] An example of the checklist for quality control on the evaluation report is in Annex 20.

- Analysis of the subject of the evaluation's framework
- Reconstruction of the intervention logic
- Findings of the evaluation by evaluation question
- Overall assessment
- Conclusions
- Recommendations (including, if relevant, directions for further research and evaluation). Recommendation should include an indication on: i) responsible actors for their implementation, ii) their importance and iii) their urgency.
- A list of dissemination materials produced and of dissemination initiatives performed

The Final Report is accompanied by the following annexes:

The ToR

The methodological approach

The evaluation matrix (data collection and analysis by EQs' indicators)

List of documents consulted

List of persons met

A statement of the ISG/RG and/or of the expert panel, highlighting any relevant issue and reporting minority opinions if relevant.

The evaluator transmits to DG NEAR or to the Delegation all database created, in a form that respects the confidentiality of all respondents and of all individuals whose data they have used.

*Source: DG NEAR*

**The utility of an evaluation also resides in the quality of reporting**, mainly that of the final report.

- **The report has to be an effective communication tool.** The results should be communicated clearly, accurately and appropriately. The report should be reader friendly and accessible to all type of readers and be self-contained, so to allow the reading also to people who have not followed the entire process. The Executive Summary should be maximum 5 of the total length and be able to provide the overall context of the evaluation and the main messages in terms of findings, conclusions and recommendations. The Executive Summary should be a stand-alone document.
- **Respect of the evaluation framework and rigour of the analysis** in order to fulfil the objective of the evaluation and expectations of the users. Independence, impartiality and evidence for the findings and conclusions are also crucial to ensure credibility of the work, thus making the report an instrument usable by the users for future improvements.
- **Usefulness of the recommendations.** Value of an evaluation report is also given by its capacity to incise into the processes and facilitate improvements. In order to do this, recommendations have to be clear, implementable and structured in such a way that the interested parties could take "inspiration" from it. Their quality thus impacts the follow-up processes.

Once the draft report is ready, evaluators present it during a final (debriefing) meeting.

Participants in the final (debriefing) meeting include at the minimum: the entire evaluation team, possibly including all team members, the evaluation manager, a representative of DG NEAR HQ's Monitoring & Evaluation service, the ISG/RG. Representatives of the stakeholders should also be invited to take part in the meeting. The evaluation manager leads the discussion, which focuses on:

- Findings, judgments, and conclusions
- The activities needed to satisfactorily concluding the work
- Dissemination activities and deliverables.

After the meeting, the evaluation manager organises written feedback, synthesising the discussion and integrating it with further comments if needed.

#### 13.3.1.2. Evaluation Staff Working Document

For major evaluations, undertaken by DG NEAR HQ according to the Better Regulation, an evaluation **Staff Working Document** (SWD) should be prepared. The SWD is drawn on the basis of the work done by the external evaluator and /or by the DG NEAR services. It summarises the results and conclusions of the evaluation. The mandatory format of the SWD is provided in the Better Regulation Package, toolbox "Staff working document"[219].

The Evaluation Scrutiny Board will review the draft evaluation SWD and the related inter-service consultation cannot start before the Board issues its opinion. The package to be sent to the Board should include the draft SWD, the associated executive summary, the minutes of the last ISG meeting, and where relevant the contractor's report and the associated the assessment grid.

### 13.4. Dissemination and Follow-up

#### *13.4.1. Disseminating the evaluation results*

**Evaluation is an opportunity to assess the performance of an EU policy or action and feed any lessons learned into the next round of decision making in a timely manner. As such, evaluation is one of the main components in decision-making and effectiveness of aid is strongly linked to the usefulness of evaluations**. EC/EUDs (and national partners) must draw lessons from evaluations and incorporate evaluation findings and recommendations into planning/programming, implementation and monitoring and evaluation (plans and exercises).

By conducting evaluations in a transparent manner the Commission is providing an account of its actions to all interested stakeholders and EU citizens. In this framework, dissemination of the evaluation results is a key function in an evaluation, it is part of a learning process and it **promotes the use of evaluations in the immediate framework in which the evaluation was carried out** (a relevant intervention, country, region, sector), but also in other relevant frameworks that can directly or indirectly benefit from the evaluation results. Indeed, **evaluations can also build knowledge for generalisation and wider application** in other processes and frameworks.

In the case of IPA II, results of evaluations shall be discussed by the IPA monitoring committee and the sectorial monitoring committees, as appropriate.

The following dissemination activities can be considered:

- **Dissemination results seminars** as the concluding step of an evaluation. Depending on the evaluation exercise, they can take place at EC HQ's or in partner countries. Their objectives are that of presenting and discussing with key stakeholders the results of the evaluation. The seminars should also be an occasion to stimulate a debate on some specific issues covered in the evaluation so as to provide additional inputs to EC and national policy makers and operational staff.
- **Organisation of thematic discussions** (via, among other, of *ad hoc* round tables) on specific issues, combining the different experiences gained during the evaluation. These discussions could be the occasion for identifying common recommendations and lessons learnt, thus providing strong evidence to use evaluation for influencing policy makers.

[219] European Commission, 2015b: tool 47.

Dissemination **results seminars and/or thematic discussions** could be done within:

- Top management of DG NEAR/EUD management meetings
- EU/partner countries dialogues
- IPA II Monitoring Committees
- ENI/IPA II Committees (with representatives of the EU Member States, etc.)
- Thematic centres' network meetings
- Multi-beneficiary meetings, Country days
- Cooperation days
- Etc.

Together with the evaluator, and keeping into account stakeholders' requests (if any), the evaluation manager and the ISG/RG can also identify **specific audiences** that should receive feedback (e.g., partner country authorities, minority groups, refugees, or schools) and ask for easily accessible materials.

- **Distribution** of the final approved report and other deliverables, if any, **via mail and e-mails.** Specific deliverables, other than the final report and the executive summary, can be identified depending on the target audiences. These can be evaluation briefs, brochures, bulletins or similar, aiming at highlighting specific findings of the evaluation or at synthetizing the executive summary with a non-technical language so as to reach a wider audience.
- **Publication on both Internet and EC intranet**, including its insertion into EVAL module.
- The **minimum requirements** refer to the **final evaluation report** and its **executive summary** (in English, German, and French), which must be **published** both in paper form[220] and on DG NEAR/EUDs/national authorities website, depending on who has commissioned the evaluation. It is important to introduce these deliverables with a paragraph aiming at 'selling' the product.

To maximise transparency and ease access:

- Any report must be published in the EU bookshop
- The final evaluation Staff Working Document, alongside the final evaluation roadmap, consultation strategy and related consultation documents, Terms of Reference and associated final contractors report (if applicable) and the opinion of the Regulatory Scrutiny Board (if applicable[221]) must be published centrally on Europa web page.

### *13.4.2. Follow-up action plan after completion of the evaluation*

Evaluation is not the end of the process. Completing the evaluation Staff Working Document (where relevant) and disseminating the evaluation findings should stimulate discussion and lead to the identification of appropriate follow-up actions to **put into practice the lessons learned and feed the evaluation findings into the next cycle of decision-making**. The evaluation results must feed into the

[220] It is important to carefully calculate the number of printed copies. Nowadays, except in particular contexts in which Internet access is difficult for various reasons, priority should be given to electronic versions.

[221] As per the Better Regulation, the Board examines and issues opinions on all the Commission's draft impact assessments and of major evaluations and "fitness checks" of existing legislation. This does not apply to project/programme evaluations.

Annual Activity Reports and related follow up actions must be identified in the Annual Management Plans of the Commission Services.

Follow-up can take many forms, such as: impact assessments**,** improving guidance or further monitoring, progress review meetings with the people in charge of the adoption and implementation of the proposed recommendations, etc. Identifying and sharing planned follow-up actions is part of accepting responsibility and accountability for EU actions and ensures transparency.

The preparation of the follow-up action plan[222] is a responsibility of the evaluation manager, together with the ISSG/RG. The evaluation findings and recommendations are assessed by the ISSG/RG in terms of their relevance and usefulness. The action table reflects the ISSG/RG views on the recommendations (whether they are accepted or not, and if not why) and what are the follow-up actions planned for the accepted recommendations. As regards recommendations that are not accepted, the ISSG/RG needs to argument its decision. Subsequently, follow-up measures on the accepted recommendations should be identified, pointing out the relevant units in charge of the action implementation and the timeframe for implementation.

The commonly agreed follow-up table is brought to the management for approval. It is disseminated afterwards to the DG NEAR/EUD/national authorities for information and implementation. The progress in realisation of the follow-up actions is regularly monitored and reported through the EAMR HQ and Annual Activity Report.

For the follow-up on the evaluation recommendations, DG NEAR will use the EVAL module[223], developed by DEVCO to register, monitor and report on the implementation of the evaluation recommendations. The module is accessible by DG NEAR HQ and EUDs.

**TABLE 18: SYNTHESIS OF EVALUATION PHASES AND MAIN RELATED ACTIVITIES AND DELIVERABLES**

| Phases | Activities | Deliverables (& meetings) |
|---|---|---|
| **PREPARATORY** | ▪ Evaluation design, expertise and budget | ☑ Evaluation Plan<br>☑ Road map, when relevant<br>☑ Terms of Reference (ToR), with<br>☑ first draft of Intervention logic<br>☑ first draft of Evaluation questions headlines |
| **INCEPTION:** STRUCTURING | ▪ Data collection & analysis methods<br>▪ Interviews at EC HQ (& country visits if relevant)<br>▪ EU Intervention's rationale<br>▪ Inventory of the EU actions (at thematic/country/regional levels) and analysis, if relevant<br>▪ Report writing (& quality control) | ☑ Final Intervention Logic<br>☑ Evaluation Questions, with judgment criteria & indicators<br>☑ Data analysis and collection methods<br>☑ EU actions inventory (& that of other donors if relevant)<br>☑ Inception Report<br>☑ Consultation strategy<br>☑ Meeting with ISG/RG |
| **DESK:** DATA COLLECTION & ANALYSIS | ▪ Document analysis (focused on the EQs)<br>▪ Interviews<br>▪ Methodological design (specific to Field visit)<br>▪ Selective analysis of other Donors | ☑ Desk report<br>☑ Field visit methodology<br>☑ Meeting with ISG/RG |

[222] Refer to Annex 21.

[223] Refer to https://myintracomm.ec.europa.eu/dg/devco/resources-procedures/it-information-systems/eval/Pages/index.aspx.

| | | |
|---|---|---|
| | actions, if relevant<br>▪ Report writing (& quality control) | |
| **FIELD** | | |
| *(Plans, methodology and budgets for the field phase are outlined and agreed upon, all along the previous phases, since the preparatory one)* | ▪ Initial meeting at country level<br>▪ Data collection and analysis<br>▪ Note writing on field phase findings<br>▪ Discussion of the findings of the Field Phase with EC/EUD & national counterparts | ☑ Briefing & debriefing with EUDs at country level<br>☑ Country/regional/thematic Note, if relevant<br>☑ Debriefing with ISG/RG |
| ***SYNTHESIS*** | ▪ Expressing findings (focus on the EQs)<br>▪ Overall assessment, Conclusions and Recommendations<br>▪ Synthesis report writing (& quality control) | ☑ Matrix of EQs, judgement criteria, indicators & analysis<br>☑ Synthesis report<br>☑ Meeting with ISG/RG |
| ***DISSEMINATION AND FOLLOW UP*** | ▪ Staff Working Document writing<br>▪ Seminar, roundtables, etc.<br>▪ Briefs, brochures, bulletins, etc. writing<br>▪ Action plan writing | ☑ Staff Working Document<br>☑ Briefs, brochures, bulletins, etc. if relevant<br>☑ Action plan |

*Source: DG NEAR*

# REFERENCES

Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions on "Strengthening the foundations of Smart Regulation – improving evaluation (COM(2013) 686 final of 02/10/2013)

European Commission (2013), *EVALSED: The resource for the Evaluation of Socio-Economic Development*

European Commission (2015a), *Guidance Document on Evaluation Plans. Terms of Reference for Impact Evaluations Guidance on Quality Management of External Evaluation.* Commission Staff Document

European Commission (2015b), *Better Regulation Guidelines.* Commission Staff Document

European Commission (2015c) *Better Regulation Toolbox.* Commission Staff Document http://ec.europa.eu/smart-regulation/index_en.htm

European Commission (2015d) *Launching the EU International Cooperation and Development Results Framework* (2015)80 final https://ec.europa.eu/europeaid/staff-working-document-launching-eu-international-cooperation-and-development-results-framework_en

European Commission (2015e) *Commission's Forward Planning of Evaluations and Studies 2015 and Beyond* http://ec.europa.eu/smart-regulation/evaluation/docs/evaluation_forward_plan_2015_en.pdf

European Commission (2015f) 182 final - "Gender Equality and Women's Empowerment: Transforming the Lives of Girls and Women through EU External Relations 2016-2020, Commission Staff Document

Evaluating EU activities. a practical guide for the commission services, DG Budget Evaluation Unit, July 2004

Framework Agreement IPA II, where it is set out the link between monitoring system and evaluation under IPA II;

GAO-PEMD (1990) Prospective Evaluation Methods. The Prospective Evaluation Synthesis. Washington, D.C.: United States General Accounting Office

Handbook on planning, monitoring and evaluating for development results, United Nations Development Programme, 2009

Internal Monitoring Guidelines, DG DEVCO, 2007 : https://ec.europa.eu/europeaid/sites/devco/files/methodology-tools-and-methods-series-strengthening-project-internal-monitoring-200706_en_2.pdf

Jody Zall Kusek & Ray C. Rist, Ten Steps to a Results-Based Monitoring and Evaluation System – A Handbook for development practitioners, Global HIV/AIDS Program and Operations Evaluation Department, the World Bank, 2004

Methodological basis for evaluation, Joint Evaluation Unit, DG DEVCO, 2006

Morris M. (2008), *Evaluation Ethics for Best practice. Cases and Commentaries*, the Guilford Press, New York

Note on "The role of evaluation and of the Evaluation sector in DG ELARG" (ARES(2014)2501607 of 29/07/2014) which puts forward the main principles and modus operandi of evaluation in DG Enlargement;

Note on securing budget for project/programme evaluations ARES ELARG/A3/OC/D(2014)4080381

Note on the Monitoring and Evaluation functions (DG NEAR as of 2 June 2015)

Note on IPA II performance monitoring , ARES note n. 3842820, dated 22/07/2016

O'Connell, D.; Raison, J.; Hatfield-Dodds, S.; Braid, A.; Cowie, A.; Littleboy, A.; Wiedmann, T.; Clark, M. (2013) *Designing for Action: Principles of Effective Sustainability Measurement* Commonwealth Scientific and Industrial Research Organization for the World Economic Forum Global Agenda Council on Measuring Sustainability

OECD DAC Network on Development Evaluation (2010b) *Evaluating Development Co-Operation. Summary of Key Norms and Standards*.

OECD DAC Working Party on Aid Evaluation (2010a) *Glossary of Key Terms in Evaluation and Results Based Management*

Pawson, R. (2006) *Evidence-based Policy. A Realist Perspective*. London: Sage.

Picciotto, R. (2013), "Evaluation Independence in Organizations", in *Journal of Multidisciplinary evaluation*, 9 (20), pp. 18-31, www.jmde.com

Regulation (EU) No 1268/2012 of 29 October 2012 on the rules of application of Regulation (EU, Euratom) No 966/2012 of the European Parliament and of the Council on the financial rules applicable to the general budget of the Union

Regulation (EU) No 231/2014 of the European Parliament and of the Council of 11 March 2014 establishing an Instrument for Pre-accession Assistance (IPA II)

Regulation (EU) No 232/2014 of the European Parliament and of the Council of 11 March 2014 establishing a European Neighbourhood Instrument

Regulation (EU) No 236/2014 of the European Parliament and of the Council of 11 March 2014 laying down common rules and procedures for the implementation of the Union's instruments for financing external action

Regulation (EU) No 447/2014 of 2 May 2014 on the specific rules for implementing Regulation (EU) No 231/2014 of the European Parliament and of the Council establishing an Instrument for Pre-accession assistance (IPA II)

Regulation (EU, EURATOM) No 966/2012 of the European Parliament and of the Council of 25 October 2012 on the financial rules applicable to the general budget of the Union and repealing Council Regulation (EC, Euratom) No 1605/2002

ROM Guidelines, DG DEVCO, 2015: https://myintracomm.ec.europa.eu/dg/devco/quality-impact/monitoring-rom/Pages/external-monitoring.aspx)

Rønsholt, F.E. (2014) *Review of Budget Support Evaluations*. Evaluation Studies. DANIDA

Stame, N. and L. Leone (2014) *Attività di supporto conoscitivo per lo sviluppo di un codice etico dei Nuclei di valutazione e verifica degli investimenti pubblici*.

Van der Linde, M. and Valmarana, C. (2011) Evaluating Budget Support. Methodological Approach. OECD-DAC