Guide for Drafting the Evaluation Plans of the 2021-2027 Cohesion Policy in Romania

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

Guide for Drafting the Evaluation Plans of the 2021-2027 Cohesion Policy in Romania[1]

DISCLAIMER

This guide is a product of the International Bank for Reconstruction and Development/the World Bank. The findings, interpretation, and conclusions expressed in this paper do not necessarily reflect the views of the Executive Directors of the World Bank or the Governments they represent. The World Bank does not guarantee the accuracy of data included in this work.
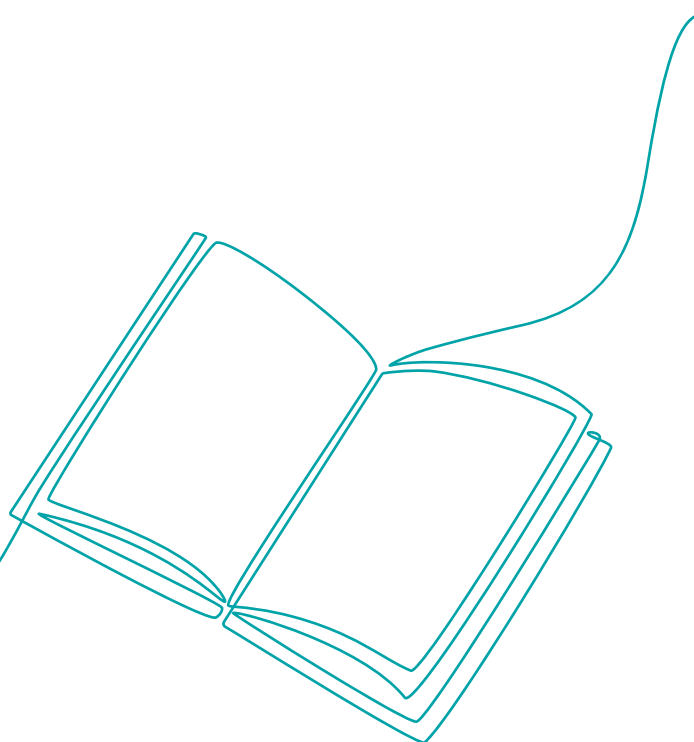
This report does not necessarily represent the position of the European Union or the Romanian Government.

COPYRIGHT STATEMENT

GLOSSARY

AMIF	Asylum, Migration and Integration Fund
BMVI	Border Management and Visa Instrument
CBA	Cost-benefit analysis
CCMAP	Coordination Committee for the Management of the Partnership Agreement
CIE	Counterfactual impact evaluation
CPR	Common Provision Regulation
DG EMPL	Directorate General for Employment, Social Affairs and Inclusion
DG REGIO	Directorate General for Regional Policy
EC	European Commission
ECU	Evaluation Central Unit
EP	Evaluation Plan
ERDF	European Regional Development Fund
ESC	Evaluation Steering Committee
ESF	European Social Fund
ESIF	European Structural and Investment Funds

EU	European Union
EU-SILC	EU Statistics on Income and Living Conditions
GDPR	General Data Protection Regulation (EU) 2016/679
IB	Intermediate body/ies
ISF	Internal Security Fund
LFS	Labour Force Survey
MA	Managing Authority/ies
MC	Monitoring Committee
MEIP	Ministry of European Investments and Projects
MS	Member State/s
NGOs	Non-governmental Organizations
OP	Operational Program/s
SC	Scientific Committee
SMEs	Small and medium sized enterprises
w.d.	Working days

INDEX

INDEX OF TABLES

INDEX OF FIGURES

INTRODUCTION

This guide aims at supporting Romanian Managing Authorities (MAs) and other key actors in the preparation of the Evaluation Plans (EPs) required by the EU Regulation 1060/2021. This guide has been prepared under the World Bank Reimbursable Advisory Services (RAS) Agreement with Romania and aimed at improving Monitoring and Evaluation (M&E) capacity in the context of EU-funded Operational Programmes (OPs). The guide is part of Output 2(b) draft "Indicators table, related metadata and evaluation plan" of the abovementioned RAS, which will be delivered officially in September 2022, as per the RAS provisions.

In Romania, the 2021-2027 programming period is characterized by increasing decentralization, which involves a high number of administrations with different levels of experiences and know-how in evaluation. The level of knowledge of M&E among institutions ranges across administrations. The guide will provide a step-by-step guide on how to design and plan evaluations, as well as guidance on how ensure compliance with the regulatory requirements in a timely and effective matter.

At the same time, the new EU regulations for Cohesion Policy provide a general framework for conducting evaluation, but do not detail what evaluations must be carried out. In comparison to the previous 2014-2020 period, when specific rules on impact evaluations were provided, Member States (MSs) are now requested to define their own evaluation policy with a high degree of freedom.

Due to the incoming new management of the operational programs and the aforementioned need for a more detailed and independent evaluation policy, the evaluation plan will be a crucial part in ensuring the execution of a sound policy which requires many participants to make this happen. The guide intends to support administrations and stakeholders in defining their evaluation plans and in promoting a coordinated evaluation strategy at the local and national level.

The guide is structured in 5 main sections:

- The **first section** explains the main definitions and concepts of evaluation and evaluation plans, in general and in the context of EU European Structural Investment Fund (ESIF).

- The **second section** ("What is necessary to know") provides a brief review of the main terms and concepts that are necessary for an evaluation plan.

- The **third section** presents the evaluation plan, as defined by DG Regio2; this document represents the objective of the officials who will use this guide.

- The **fourth section** focuses on the identification and initial design of the evaluations that will be included in the evaluation plan.

- The **fifth section** discusses the necessary elements that shape the evaluation framework that is used to create evaluation plan ; in other words, it details the governance of the evaluation processes of the operational program.

- **Annexes** provide detailed examples, templates to develop evaluation activities and other practical tools which complement this guide.

For those who are interested only in a specific part of the guide, the next table shows the main cross-references between the template of the evaluation plan and the sections of the guide.

Section of the evaluation plan	Reference to the guide
INTRODUCTION	(no specific reference)
PART 1 EVALUATION FRAMEWORK	
1.1 National governance of the evaluation	section 5.1.1
1.2 Coordination and ORGANIZATION of the evaluation at the OP level	section 5.1.2
1.3 Involvement of the stakeholders in the evaluation plan	section 5.2
1.4 Strategic approach to the evaluation	section 4
1.5 Evaluation implementation and quality	section 5.3
1.6 Dissemination and use of the findings	section 5.4
1.7 Evaluation capacity building	section 5.5
1.8 Available resources and timing of the evaluation	section 4.5
PART 2 PLAN OF THE EVALUATIONS	
2.1 Lists and timetable of the evaluations	section 4
2.2 Fiches of the planned evaluations	section 4

---

2   European Commission, 2021, *Performance, monitoring and evaluation of the European Regional Development Fund, the Cohesion Fund and the Just Transition Fund in 2021-2027*, Commission Staff Working Document.
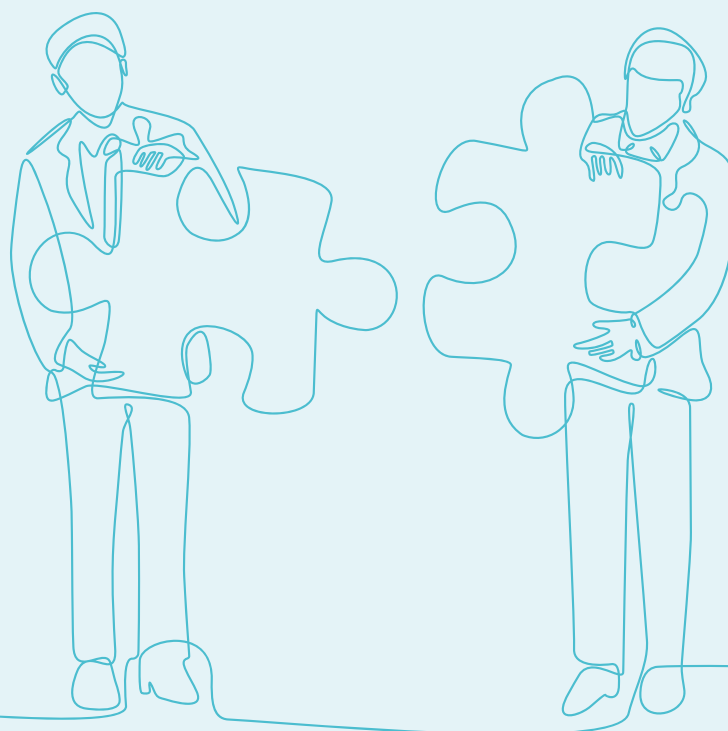
EVALUATION AND EVALUATION PLAN

1

1.1. WHAT IS EVALUATION AND WHY WE USE IT

Definition of Evaluation

Different definitions of evaluation are possible, but the ESIF evaluation is precisely described in Rossi et al (1999):

"Program evaluation is the use of social research procedures to systematically investigate the effectiveness of social intervention programs. More specifically, evaluation researchers (evaluators) use social research methods to study, appraise, and help improve social programmes in all their important aspects, including the diagnosis of the social problems they address, their conceptualization and design, their implementation and administration, their outcomes, and their efficiency."

In addition to this broad definition, it is important to underline that evaluation research differs from other social research as it aims to provide a "value" or a judgment of the programme (effective/ineffective, efficient/inefficient, etc.). In the management of a programme, other types of research like exploratory studies, literature reviews, analysis of monitoring data, etc., can be very useful in addressing operational choices, but they are not evaluations and by themselves are incapable to assess the quality of the programme.

Why evaluation is used

The reasons why evaluation is used may differ according to the context and the objectives of the evaluation. Generally, one or more of the following reasons lead to carrying out the evaluation:

- *Accountability* – where the intention is to give an account to sponsors, policy makers or taxpayers of the achievements of a programme;

- *Development* – where the intention is to improve the delivery or management of a programme or across programmes through better coordination of support;

- *Knowledge* – where the intention is to develop new knowledge and understanding of the public policy ('what works and why') to better allocate resources among policy instruments.
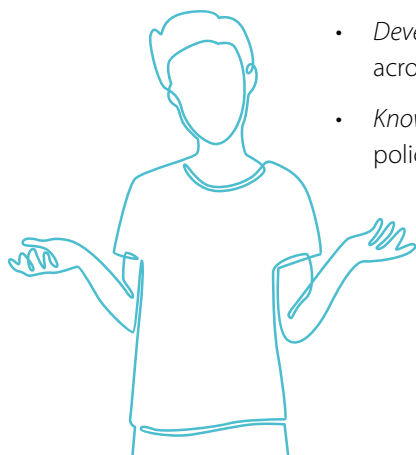
In the framework of the ESIF, all these reasons are relevant, but this does not mean that they necessarily need to coexist in a single evaluation or address all the evaluations of a programme. In addition, the three reasons trigger different approaches: an evaluation needed for *accountability* requires a computation of interventions and outputs; the *development of the programme* requires the investigation the implementation processes and governance; and for *the purpose of acquiring new knowledge*, there should be an analysis of the impact and the way it takes place.

The different types of evaluation

Various types of evaluation are used, depending on the aim of the evaluation. In general, evaluations are divided into two main groups:

- *Formative evaluations*, aimed to introduce improvements in the programme;

- *Summative evaluations*, aimed to inform decisions about whether to start, continue, expand, or stop an intervention.

The focus of the evaluation can also differ for each of these two types of evaluation (Table 1), allowing for:

- Process evaluations focused on the programme implementation; and

- Impact evaluations focused on the outcomes and effects of the programme.

TABLE 1 Main types of evaluation

	Formative evaluation	Summative evaluation
Process evaluation (focused on processes)	Intended to inform decisions about improving primarily the programme implementation	Intended to inform decisions about stop/go the implementation system of the programme or some of its interventions
Impact evaluation (focused on impact)	Intended to inform decisions about improving primarily the programme design characteristics	Intended to inform decisions about stop/go the programme or some of its interventions

From: https://www.betterevaluation.org/

"Helpdesk evaluation", the service to support the evaluation of European Regional Development Fund (ERDF) and European Social Fund (ESF) programmes funded by Directorate General for Regional Policy (DG Regio) and Directorate General for Employment, Social Affairs and Inclusion (DG Employment), has collected, and examined all the evaluations carried out in the 2014-2020 period[3]. Its repository of ERDF and ESF evaluations use the following definitions:

- *Monitoring/progress oriented*, evaluations focused on the monitoring of the interventions and aimed at analysing the progress of the interventions.

---

3   See https://ec.europa.eu/regional_policy/it/policy/evaluations/member-states/.

- *Process/implementation oriented*, evaluations focused on the implementation and management processes of the interventions.

- *Impact/results oriented*, evaluations focused on the analysis of effects and impacts of the interventions.

- *Summary of evaluations/meta-analysis*, evaluations based on synthesis of results of other and previous evaluations.

- *General studies*, the research included in the evaluation plans are connected to the evaluations but have an informative aim as opposed to an evaluative aim.

In conclusion, several types of evaluation are commonly used in ESIF and differ from one another in terms of both scope and objective. This means that the evaluation plan uses different types of evaluations, that are selected depending on which objectives are pursued and whether the focus is on implementation or impact. An evaluation plan cannot consist only of descriptive or monitoring evidence, but also needs to include significant evaluation research.

Evaluation in the decision-making process

Evaluation supports policy makers in taking the most informed decisions on public policies. Evaluation also helps designing new programmes (ex-ante evaluations), improving the implementation of the programmes (on-going evaluations) and changing existing programmes or shaping future programmes (intermediate or ex-post evaluations).

As required by Common Provision Regulation (CPR) 1060/2021 (the common ESIF guidance for the 2021-2027 period) evaluation has *"to improve the quality of the design and the implementation of the programmes"* and assess their impact (see Box 1). Therefore, while evaluation provides technical support and informs the decision-making process, it is also an obligation that MSs and MAs must respect to account for their expenditure to the European Commission (EC), other stakeholders, and taxpayers.

Evaluation is also a part of a wider system of control and quality assurance of ESIF. Other instruments such as audit, performance monitoring and surveillance of the Monitoring Committee support the correct implementation and the success of the ESIF. However, each of these instruments has its own aim and competence and are not to be used interchangeably. For instance, monitoring measures the performance of a programme,

but it does not assess its quality and effectiveness as evaluation does. An audit verifies the compliance of an implementation system with the existing rules but does not appraise the influence of the implementation on the final effects as evaluation does. As audit and monitoring cannot be confused with evaluation, evaluation cannot be used for audit or monitoring. These different instruments all contribute to the effective management of ESIF and reciprocally integrate their findings, but each of them covers a specific area of investigation and pursues different objectives.

Limits of the evaluation and frequent pitfalls

Evaluation must not be interpreted as an infallible and all-powerful instrument: such as, a kind of unquestionable judgement deriving from mathematical calculations.

As with other social sciences, evaluation suffers from many limitations due to the complexity of the evaluated phenomena which may involve the analysis of the social and economic behaviours of a great number of different subjects, but also due to the limitations in data available to measure such phenomena. To overcome these shortcomings, it is necessary that evaluation responds to scientific criteria, is aware and transparent of its limitations and focuses on policy interventions for which it can provide solid interpretations.

In this respect, ESIF evaluations must avoid some frequent pitfalls such as attempting to answer an unrealistic number of questions within the same research topic, adopting a scope of policy interventions that is inconsistent with the objectives of the analysis, using inappropriate methodologies or, adopting methodologies without the right data to implement them. Such pitfalls, among others that can occur during evaluations, can produce weak evidence and biased judgments, and can imply a loss in the authority and credibility of the evaluation.

To prevent pitfalls, the evaluation of the ESIF must adopt scientific approaches and pay attention to data availability and all other factors limiting the feasibility of the research. The careful preparation and planning of the evaluation of the ESIF programmes is an unquestionable opportunity to prevent drawbacks and reinforce the capacity of administrations responsible for the evaluations.
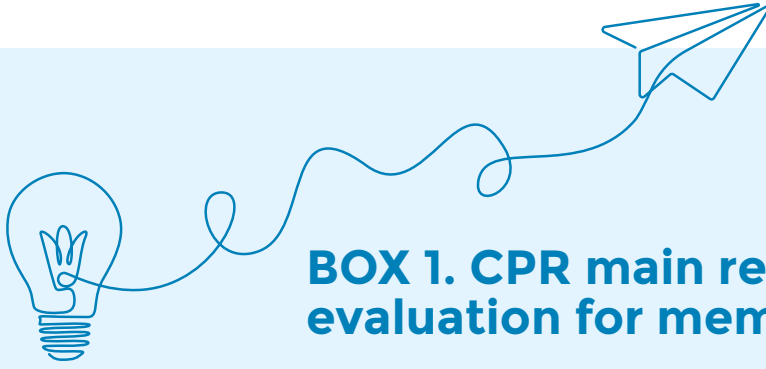
BOX 1. CPR main requirements on evaluation for member states

Common Provision Regulation 1060/2021

Article 44 - Evaluations by the Member State

1. The Member State or the managing authority shall carry out evaluations of the programmes related to one or more of the following criteria: effectiveness, efficiency, relevance, coherence and Union added value, with the aim to improve the quality of the design and implementation of programmes. Evaluations may also cover other relevant criteria, such as inclusiveness, non-discrimination and visibility, and may cover more than one programme.

2. In addition, an evaluation for each programme to assess its impact shall be carried out by June 30, 2029.

3. Evaluations shall be entrusted to internal or external experts who are functionally independent.

4. The Member State or the managing authority shall ensure the necessary procedures are set up to produce and collect the data necessary for evaluations.

5. The Member State or the managing authority shall draw up an evaluation plan which may cover more than one programme. For the AMIF, the ISF and the BMVI, that plan shall include a mid-term evaluation to be completed by March 31, 2024.

6. The Member State or the managing authority shall submit the evaluation plan to the monitoring committee no later than one year after the decision approving the programme.

7. All evaluations shall be published on the website referred to in Article 49(1).

Article 40 – Functions of the Monitoring Committee (in relation to evaluation)

1(e). The Monitoring Committee shall examine the progress made in carrying out evaluations, syntheses of evaluations and any follow-up given to findings.

2(c). The Monitoring Committee shall approve the evaluation plan and any amendment thereto.

Article 18 – Mid-term review and flexibility amount

2. The Member State shall submit an assessment for each programme on the outcome of the mid-term review, including a proposal for the definitive allocation of the flexibility amount referred to the second paragraph of Article 86(1), to the Commission by 31 March 2025.

1.2. WHY AN EVALUATION PLAN

The obligation to have an evaluation plan is stated in the regulatory requirements (see above), but is also justified by practical reasons and the purpose of delivering high quality and effective evaluations. An evaluation plan helps clarify the direction the evaluation should take and identifies the priorities and appropriate resources, time, and skills needed to accomplish the set goals.

As per the regulatory requirements stated in the EC regulations in force[4], the Member State or the managing authority would need to draw up an evaluation plan, which may cover more than one programme (Art 44 (5)). This should be submitted to the monitoring committee (MC) no later than one year after approving the programme.

For the current period 2021-2027, evaluations are required at both programme and fund level.

- The evaluations at programme level should be carried out by the Member State and assess the programmes' effectiveness, efficiency, relevance, coherence and EU added value, with the aim to improve the quality of the design and implementation of programmes (Art 44, CPR).
    - In 2024 (mid-term review) and, respectively, in 2031, the EC will examine the same criteria, this time at fund level (Art. 45, CPR).

Impact evaluations are expressly required by 2029 for the entire programme, while ex-ante evaluations are no longer compulsory, with the exception of the ex-ante assessments required for financial instruments. For the final impact evaluation, the responsible authorities can consider the entire OP or review just part of the programme that is most representative. Regardless of the approach chosen, it is compulsory to provide an impact evaluation of the programmes.

In addition, as per the provision of Art 86 (1), Member States are required to submit an assessment for each programme on the outcome of the mid-term review, including a proposal for the definitive allocation of the flexibility amount referred to the second paragraph of Article 86(1), to the Commission by 31 March 2025. While this is not an evaluation per se, the technical inputs for the required assessment could come as a result of the evaluations planned and should be planned accordingly.

While the regulations provide for simplified evaluation requirements in comparison to the 2014-

2020 programming period, the absence of detailed requirements places a higher responsibility on the shoulders of the Member States and MAs to improve the design and scope of evaluations.

Having broader regulations impacts how MS and MAs define their evaluation strategies. MSs and MAs have now increased freedom in defining their own evaluation strategy. In this context, it is important to underline that the evaluation criteria (effectiveness, efficiency, etc) have not to be necessarily assessed in the same evaluations, but can be examined separately in different evaluations. Similarly, the "evaluations of the programmes" provided by article 40 of the CPR does not mean that the entire programme has to be assessed in each evaluation: MSs can continue to implement thematic evaluations and evaluations focused on specific components or measures of the programme according to their needs and the most suitable methodological arrangements. On the other hand, all the programmes must be evaluated and appraised on their functioning and effects and the role of evaluation in supporting policy and decision making should be strengthened.

From an operational point of view, it is important to have a well-designed evaluation plan that is approved by the Monitoring Committee and ensure that evaluations run smoothly according to an adequate timetable. It also helps foster transparency and ensure that stakeholders are on the same page in regards to the purpose and use of the evaluation results. The evaluation plan is a key tool in better preparing and directing the entire implementation of the evaluation and the use of its results.

In short, the evaluation plan helps in defining evaluation priorities, scheduling evaluation activities promoted by the MA and guiding the relevant stakeholders through each step of the evaluation process. The evaluation plan also executes the following functions:

- Creates a shared understanding of the purpose(s) and its use, by users of the evaluation results;

- Fosters program transparency to stakeholders and decision-makers;

- Helps prioritize evaluation themes and questions, in line with the identified evaluation needs;

- Guides stakeholders through each step of the evaluation process by identifying the best methods and strategies, clarifying methodological aspects and

---

4   Reg no. 1060/2021 on the Common Provisions Regulations.

information needed for certain evaluations, as well as expected outcomes;

- Increases buy-in and acceptance of methods;

- Favours an early identification of the data needed to implement planned evaluation and promotes arrangements with data owners, where necessary;

- Helps come up with a realistic timeline for evaluations, in line with the needs for information to support decision-making;

- Connects multiple evaluation activities;

- Serves as an advocacy tool for evaluation resources based on negotiated priorities and established stakeholder and decision maker information needs;

- Supports evaluation capacity building among administrative officials, market partners and stakeholders;

- Provides a multi-year comprehensive document that makes explicit main activities from involvement of the stakeholders to dissemination and use of results;

- Promotes good evaluation practice.
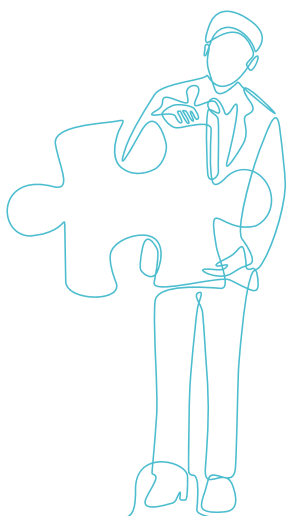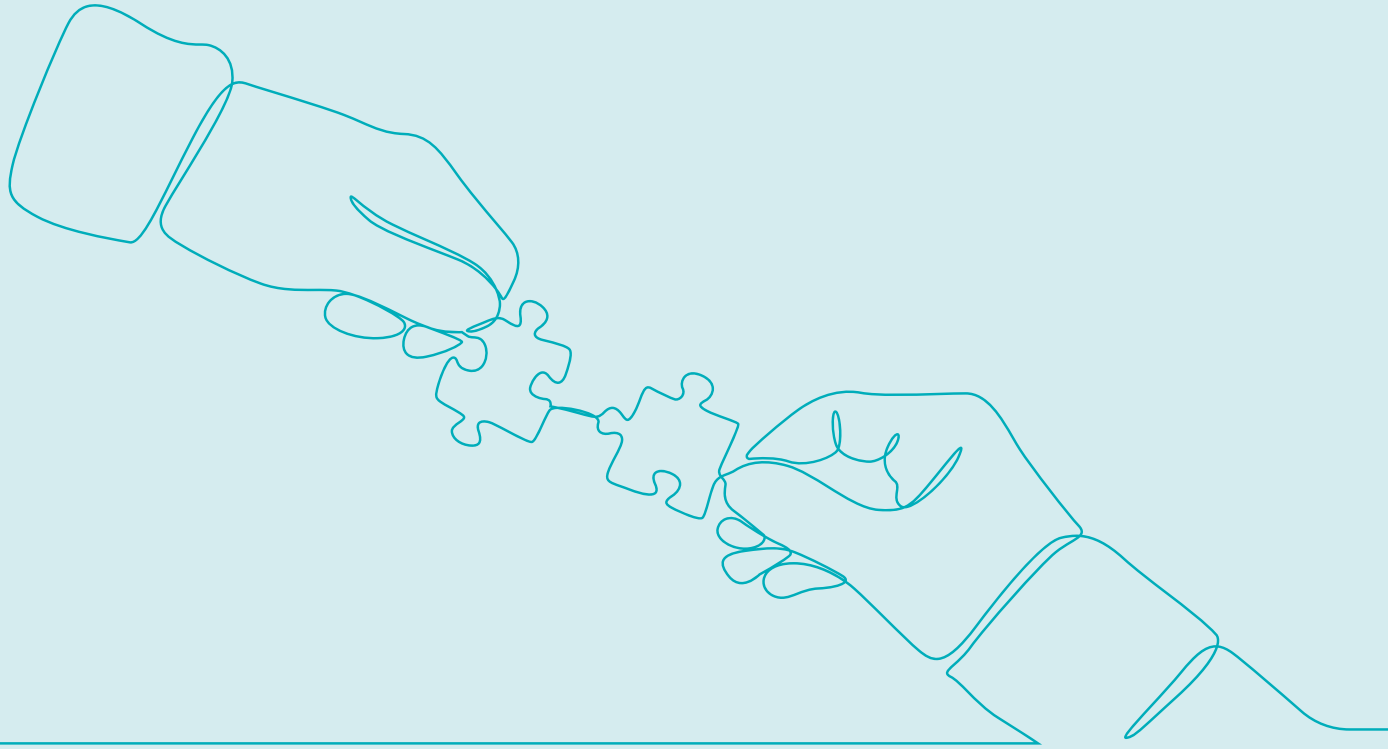
WHAT IS NECESSARY TO KNOW (IN PILLS)

2

This section presents some basic elements that are necessary to know before planning evaluations. They relate to evaluation criteria, the stakeholders, the theory of change (ToC) or programme theory, the different types of evaluation and the main methods used in evaluation. These topics play an important role in designing and planning evaluation and for this reason should be mentioned in advance. A detailed study of these topics would require much more space than available here, but this guide aims to set up the evaluation plans and a more in-depth examination of these topics will be possible through the references provided in Annex D.

2.1. THE EVALUATION CRITERIA

As mentioned in the introductory section, evaluation research must provide a judgment about the programme or the policy that is under assessment. To formulate a judgment of the performance of a programme, it is necessary to link the assessment with clear and shared criteria.

CPR indicates five evaluation criteria, which are aligned with the criteria used in the Better Regulation Guidelines[5] of the EC. Thus, it primarily refers to effectiveness, efficiency, relevance, coherence and EU added value. The first four criteria are the most common and generally used in evaluation, while the fifth (Union added value) is particularly relevant to EC policies, which needs to be assessed in comparison to other national or regional policies. Overall, these five criteria provide a basic set of principles to assess programmes.

The interpretation and the use of these criteria should be contextualized in relation to the programme and the policy measures under assessments. In general, the criteria are as follows:

Effectiveness: Effectiveness analysis considers how successful EU action has been in achieving or progressing towards its objectives. The evaluation forms an opinion on the progress made to date and the role of the intervention in delivering the observed changes. Success factors are also identified, if applicable. If the report indicates the program is not meeting expectations or failed objectives entirely, an assessment is carried out to understand the extent to which progress has fallen short of the target and what factors have influenced the failure or delay in achieving expected results.

Efficiency: Efficiency analysis examines the relationship between the resources an intervention uses and the changes generated by the respective intervention. 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 in implementation methods or policy instruments would have achieved the same benefits at less cost (or greater benefits at the same cost).

Efficiency analysis involves an examination of administrative costs and changes in outcomes as a result of implementation of ESIF interventions. It looks at how the interventions can be simplified and streamlined. Assessing efficiency in cohesion policy programmes is particularly complicated at a programmatic level, because of the variety of interventions supported under the same programme.

Relevance: Relevance looks at the relationship between the needs and problems in society and the objectives of the intervention and touches on aspects of strategy and policy design. As circumstances and context change over time, certain objectives may already have been met or, by contrast, may be no longer be relevant. Relevance analysis is important because if an intervention does not help to address current needs or problems, it is no longer appropriate. The relevance analysis coincides and is bolstered by the analysis of EU added value (see below), which looks at whether an action continues to be justified at the EU level.

Coherence: The evaluation of coherence examines the compatibility of certain paired actions. Even small changes in the way one intervention is designed or implemented can trigger improvements or inconsistencies with other ongoing actions. The comparison points that are used to assess coherence may vary according to the timing and to the level of coherence being considered.

- Checking for *internal coherence* involves looking at how the various components of a programme operate together to achieve the programme objectives (we can have coherence between measures in a policy or specific objective, coherence between policies or specific objectives).

- Similarly checking for *external coherence* can entail, for example, looking at the coherence between interventions within the same policy field supported by different EU or national instruments (e.g., coherence between national and regional programmes or with non-cohesion instruments) or between different areas, which may have to work together (e.g., Small and medium sized enterprises (SMEs) interventions and environmental policy).

EU added value: The assessment of the EU added value looks for changes that can reasonably be attributed to the EU intervention, in addition to what is typically expected from national actions by the Member States. Different approaches can be relevant to evaluating EU added value; added value may be demonstrated by showing:

- The role EU financing plays in funding actions that would not otherwise take place, that take place on a greater scale or that happen earlier than would otherwise be the case;

---

5   The five evaluation criteria referred to in the regulation, aligned with the Better Regulation Guidelines https://ec.europa.eu/info/sites/info/files/file_import/better-regulation-toolbox-47_en_0.pdf

- The extent to which the method (design and practices) of cohesion policy programming brings benefits in relation to other national practices;

- How the interventions under the cohesion policy programmes contribute to EU strategic goals and priorities;

- If the cohesion policy programmes introduce new strategic and innovative policies in the national policy environment.

In interventions where the Cohesion Policy is the main source of investment in the programme area, the answer to the question of EU added value may simply involve the simple reiteration of the reasons why the interventions are financed from the cohesion policy or refer to the efficiency and effectiveness analysis carried out. In theory, a strict analysis of economic added value would look to see if interventions supported bring higher returns beyond what could reasonably have been expected from investments financed by national authorities. In practice, this is often challenging in shared management, as practically all interventions involve both EU and national financing.

Article 44 of the CPR, which includes provisions on the evaluation under the responsibility of the Member States, indicates that three other criteria may also be used. These criteria are less common in evaluation but may be pertinent in some cases when there is a need to assess some general or horizontal objectives of the ESIF investments. These additional criteria are:

- **Inclusiveness**, which is the capacity of the programmes to include and assist different segments of population and especially disadvantaged populations lacking in public support (e.g., people in social and economic marginalization, convicted people, etc.). This criterion is mainly concerning social policies, but it could be also applied to SME interventions, when looking at the most vulnerable enterprises or those most difficult to assist.

- **Non-discrimination**, which is the extent to which all the individuals – or the individual enterprises – have an equal and fair chance to access opportunities made available by the programme. This criterion is particularly relevant to population victims of negative stereotypes (e.g., Roma, LGBT communities, ethnic or religious minorities, etc.).

- **Visibility**, which looks at how the communication activities of the programme make the EU policy visible to the interested population and aims to appraise the public awareness of the EU financial and policy effort.

The use of the evaluation criteria is not automated and must be correctly selected and contextualized to each programme and intervention. To this aim, some considerations on evaluation criteria are necessary:

- *Evaluation criteria are not evaluation questions*. Criteria are needed to formulate an appraisal, while evaluation questions explain the intended goal of the evaluation; hence, "effectiveness" or "efficiency" cannot be the questions on whom an evaluation can be designed. However, the evaluation questions need some criteria to verify the "value" and the quality of the programme; in this respect, many of the usual evaluation questions can be referred to in the initial five criteria (see Box 2).

- *Each criterion must be contextualized*. The above definitions of the criteria refer to their conceptual interpretation. However, in the evaluation, the criteria should be specified in relation to the policy context under examination (objectives, policy instruments, involved actors, socio-economic conditions, etc.). For instance, efficiency may be measured in terms of unit cost of output, time spent to implement measures, or quality of the implementation; depending on what is requested, one or all three approaches to an efficiency assessment could be considered in an evaluation.

- *Evaluation criteria can produce conflicting results*. Trade-offs between results of the criteria may occur and are fairly logical; for instance, an intervention may be effective, as it is capable of reaching its objectives, but it also may be inefficient in terms of time or resources spent or in its implementation. In many cases, this or other possible "conflicts" are not surprising, but they must be explained, since they make the general assessment of the program and the related policy decision more difficult.

- *Evaluation criteria are not limited to the five criteria mentioned above (effectiveness, efficiency, relevance, coherence, EU added value)*. The five criteria mentioned are the most used in evaluation, but other criteria such as innovativeness, participation and partnership might be used according to the objectives of the evaluation; for instance, *innovativeness* can be assessed to verify if the program has promoted new policies and effective instruments. *Participation* can be used to examine to what extent some measures were implemented and supported by their main stakeholders. Finally, partnership can also be assessed, to appraise the capacity of large partnerships to work together and to contribute to higher quality policy solutions.
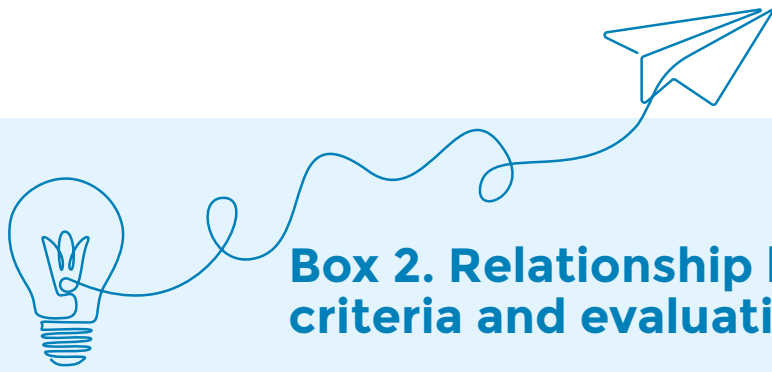
Box 2. Relationship between evaluation criteria and evaluation questions

Typical examples of effectiveness questions

- What have been the (quantitative and qualitative) effects of the intervention?
- To what extent do the observed effects link to the intervention?
- To what extent can these changes/effects be credited to the intervention?
- To what extent can factors influencing the observed achievements be linked to the EU intervention?
- For spending programs, did the associated EU anti-fraud measures allow for the prevention and timely detection of fraud?

Typical examples of efficiency questions

- To what extent has the intervention been cost effective?
- To what extent are the costs of the intervention justified, given the changes/effects it has achieved?
- To what extent are the costs associated with the intervention proportionate to the benefits it has generated? What factors are influencing any discrepancies? How do these factors link to the intervention?
- To what extent do factors linked to the intervention influence the efficiency with which the observed achievements were attained? What other factors influence the costs and benefits?
- How proportionate were the costs of the intervention borne by different stakeholder groups, taking into account the distribution of associated benefits?
- If there are significant differences in costs (or benefits) between Member States, what is causing them? How do these differences link to the intervention?
- How timely and efficient is the intervention's process for reporting and monitoring?

In order to apply the evaluation criteria correctly, the evaluator needs to identify some robust evidence to conclude whether an intervention is effective and efficient or not. Quantitative or qualitative indicators generally make up this evidence and provide the scales for measuring the effectiveness or the efficiency of the other criteria. To this end, the evaluation methodology should identify accurate and consistent indicators to answer the evaluation questions and assess the programme according to the related criteria. The methodology must also specify quantitative data or qualitative information needed for quantifying the indicators and the way to collect them.

The assessment of the impact of the OP deserves a final consideration. This assessment is explicitly requested by the CPR (see Box 1), as well as the use of the evaluation

criteria. However, impact is not an evaluation criterion. The impact is a type of effect: generally, it refers to the difference between what happened and what would have happened without the intervention. An "impact evaluation" is a type of evaluation focused on the measurement and the explanation of an intervention's impacts. This type of evaluation answers questions such as: what difference does the intervention make? And why?

Not all interventions need impact evaluations or have to be judged according to all the above criteria. The strategic character of some measures, the ability to replicate them in the future, their budget or their innovative character, and previous literature or knowledge on the subject, are factors that influence the opportunity to assess the impact of the measures.

Typical examples of relevance questions

- To what extent is the intervention still relevant?
- To what extent have the (original) objectives proven to have been appropriate for the intervention in question?
- How well do the (original) objectives of the intervention (still) correspond to the needs within the EU?
- How has the intervention adapted to subsequent technological or scientific advances? (N.B. Could include issues related to the specify policy here e.g., social, environmental or to implementation, reporting and compliance)
- How relevant is the EU intervention to EU citizens?

Typical examples of coherence questions

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

Typical questions on EU added value

- What is the additional value resulting from the EU intervention(s), compared to standard expectations from Member States acting at national and/or regional levels?
- What would be the most likely consequences of stopping or withdrawing the existing EU intervention?

Source: Better regulation toolkit of the EC (https://ec.europa.eu/info/law/law-making-process/ planning-and-proposing-law/better-regulation-why-and-how/better-regulation-guidelines- and-toolbox/better-regulation-toolbox_en)
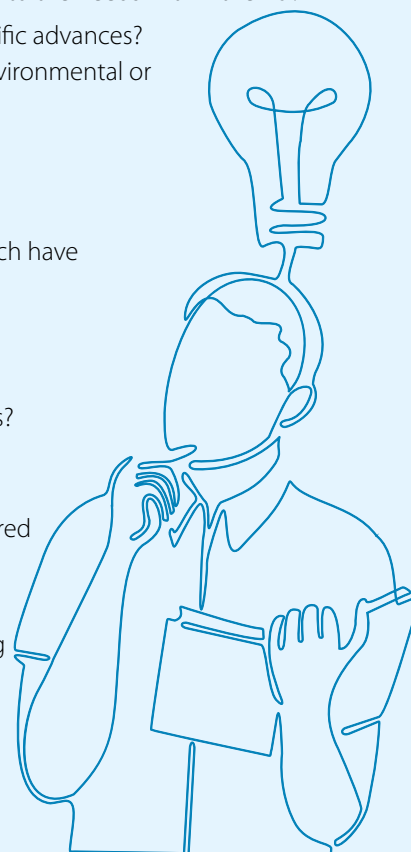
2.2. THE STAKEHOLDERS AND THEIR ROLE[6]

Stakeholders are people or organizations with a direct interest in the evaluation. They include primary intended users, implementers of the interventions and other direct and indirect beneficiaries. Understanding the objectives and interests of these stakeholders is crucial for producing useful and high quality evaluations. The stakeholders' perspective, in fact, feeds evaluation planning, communication strategies during and after the evaluation and supports the use of evaluation findings.

Stakeholders may be classified according to their role in the programme.

- Decision-makers, who will use and make decisions based on the evaluation findings are a key group of stakeholders; they include the Managing Authority (MA) and the other administration involved in the program.

- Other stakeholders include people who will be affected by decisions made during or after the evaluation (program staff, program participants and beneficiaries) and secondary users of the evaluation findings.

- Evaluation findings are also of interest to policy makers and social partners and organizations.

There is no unique strategy for involving stakeholders in the evaluation. Different stakeholders can be involved for different purposes and at different phases of evaluation planning and implementation. It is not always possible or useful to engage all potential stakeholders, and consequently, a specific strategy should be developed in each evaluation.

---

6   Partially adapted from https://www.betterevaluation.org/en/rainbow_framework/manage/understand_engage_stakeholders.

The institutional arrangements already defined propose some methods to involve some stakeholders, such as the Monitoring Committee, Evaluation Steering Committee or other key actors with a role in policy making in the area covered by the programme (as we will describe further in this Guide). In defining this strategy and selecting the stakeholders to involve, it is important to remember that the involvement of the right stakeholders during evaluation planning and implementation can improve the quality of the evaluation through the following ways:

- Offer a general view on what is considered a reliable, high quality and useful evaluation;

- Contribute to the definition of the ToC of the program and the identification of the key evaluation questions;

- Facilitate data collection;

- Help to interpret collected data;

- Reinforce the use of the evaluation's findings by building knowledge about and support for the evaluation.

Engaging stakeholders also helps to manage risks related to a controversial intervention in which key stakeholders have opposing views. It is important to comprehend different views on what will be considered a reliable result of the intervention.

To ensure that key groups of stakeholders are included at appropriate points during the evaluation, it is important to consider the following questions:

- Who are the stakeholders for this evaluation?

- What is their role in the evaluation?

- What do they need to learn from the evaluation?

- How will the findings and recommendations be communicated?

- How will the findings and recommendations be used?

- What do intended users view as credible information?

- How will the evaluation findings be communicated?

In the EU programming, stakeholders generally include the following:

- **Management staff** and several Departments and Intermediate Bodies involved in the management of the programme;

- **Other institutions** with common or relevant interest (EC, National State where Regions manage the programme);

- **Politicians** (including opposition party members), community members, and other interest groups;

- **Beneficiaries and final beneficiaries -** Those persons, groups, and organizations, who directly benefit from the intervention;

- **Indirect beneficiaries -** Those persons, groups, and organizations who have no direct contact with an intervention but who benefit from it or are somehow affected by it nonetheless;

- **Social partner organizations** (employer and trade union representatives, representatives of nongovernmental organizations and other third-party organizations)

- **Non-governmental Organizations (NGOs) and international organizations,** with activities in the areas evaluated

- **Representatives of the academia**, etc.

The involvement of the stakeholders can differ according to their relevance to the evaluated intervention. As mentioned above, three main factors can help to distinguish the relevance of the stakeholders: level of interest in the policy, influence on the design, or influence on policy implementation. Figure 1 below summaries the potential groups of stakeholders according to the three factors.

FIGURE 1 **Stakeholder Mapping Matrix**
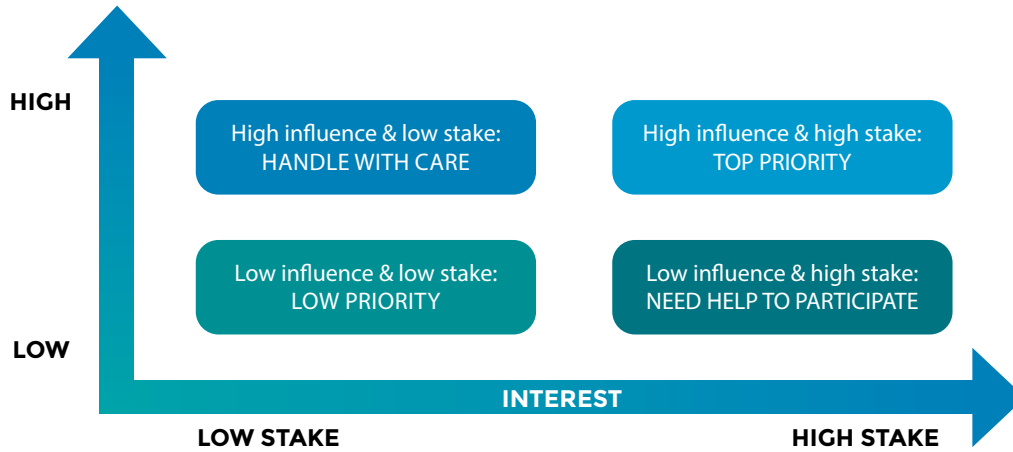
Source: EC, Better regulation Toolbox

Stakeholders may participate in focus groups to discuss the type, scope, and criteria for the evaluation. The focus groups may entail providing feedback on evaluation questions and key issues; exploring the evaluation methodology, terms of reference, work plan, and data collection strategy and process; and collecting and distributing feedback on results.

Engaging key stakeholders early in the planning stage will address any conflicts that may arise between the evaluation managers and the stakeholders and allow plenty of time to address concerns. It is important to resolve any disagreements or misunderstandings about the overall purpose or specific evaluation questions among stakeholders. The consultative process will help foster the buy-in or share ownership that will be valuable in later stages of managing the evaluation.

2.3. THE CHARACTERISTICS OF THE INTERVENTIONS TO EVALUATE: THEORY OF CHANGE

To design an evaluation or plan a set of evaluations, it is necessary to reflect on the characteristics of the programme to evaluate. These characteristics affect the definition of the evaluation questions, the methodology to use, the time when the evaluation must be carried out, and how the findings should be used.

The theory of Change (ToC) is the most effective way to summarize and visualize the main characteristics of a programme, a part of a programme or a single measure. A

ToC, or programme theory[7], explains how an intervention (a project, a programme, a policy, a strategy) contributes to a series of results that produce the intended impact. In other words, the ToC explains how the sequence of causalities is triggered - or expected to be triggered - by the programme and identifies the conditions that caused the sequence. Figure 2 shows a simple causality chain of an intervention targeting the SMEs competitiveness based on subsidies for investments.

FIGURE 2 **Causality chain for interventions SMEs competitiveness**
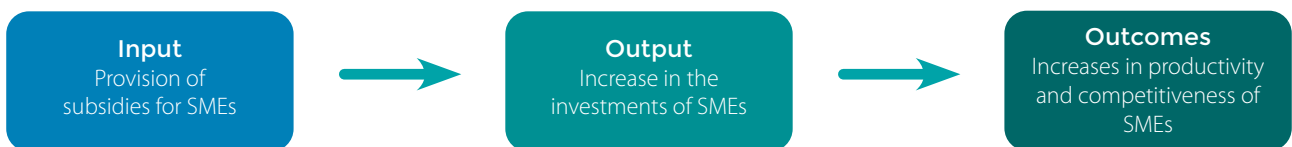
Source: Developed by Authors

---

7   In this Guide Theory of Change and program theory are used as synonymous; however, some scholars distinguish between the two. For instance, Lisa Wyatt Knowlton and Cynthia C. Phillips in their *"The Logic Model Guidebook"*, 2013, argue: "a theory of change model is simply a general representation of how you believe change will occur; a program logic model details resources, planned activities, and their outputs and outcomes over time that reflect intended results."

Figure 2 is an example of "intervention logic", which is often used to represent the multilevel set of instruments and objectives in EU programs. The ToC will build on the intervention logic. To this end, it will bring additional details and explicitly state the assumptions or the conditions that make it possible to move from one point to another in the chain. Figure 3 shows an example of a possible ToC defined according to the previous example of the subsidy for SMEs.

FIGURE 3 **Possible ToC for the SMEs subsidies (based on the causality chain presented above)**
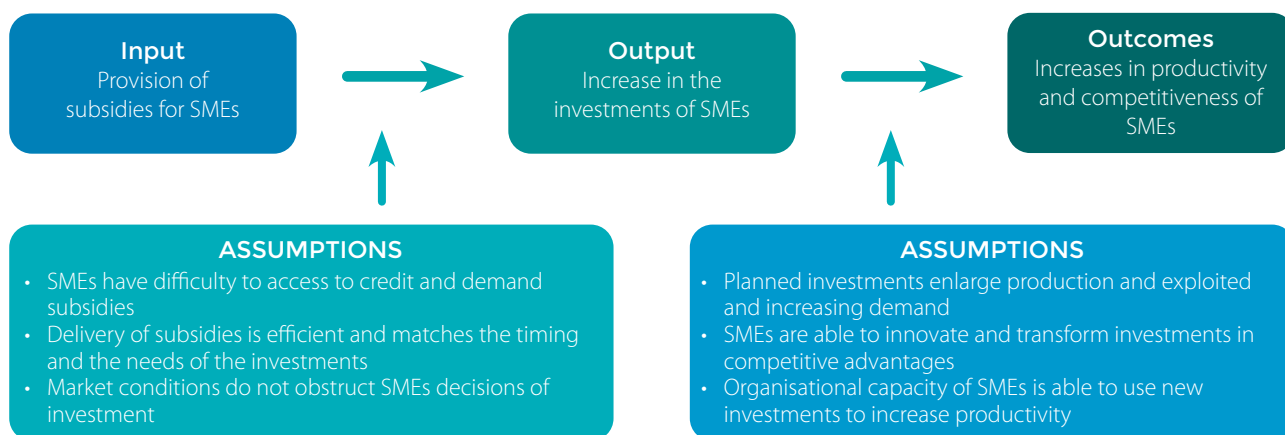
Source: Developed by Authors

The term "Assumptions" in the context of ToC may refer to general reforms or changes in the legal framework, as well as economic or social conditions, institutional arrangements, and behaviors of companies or individuals. Often these assumptions are also called "mechanisms", meaning the mechanisms that make a specific cause-effect relation possible. Assumptions are more difficult to identify because they are often implicit in the program and may depend on subjective interpretations of the program.

Two further observations are important:

- In general, the ToC tends to be more complex than that presented in the example above. Inputs and effects can be multiple, as well as synergies between them. Consequently, also the identification of the assumptions can be more difficult to articulate.

- ToC can be defined for **each of the programming levels:** project, action, priority axis, programme or Partnership Agreement. The complexity of the ToC is directly correlated with the complexity of the programming level concerned. For example, because the programme level is more complex, the ToC at this level is expected to be more detailed than a ToC at the project level. In addition, at programme level, the ToC has to focus on main relationships and assumptions underlying the programme, otherwise it is useless for analysis at this level. On the other hand, at the single action or group of action level, the ToC has to be more specific with in-depth analysis.

The ToC provides a conceptual framework that can be applied to the programme's design, monitoring, evaluation, and communication. In the context of the evaluation planning, the ToC allows to understand what programmatic components are most important to evaluate, what should be the key questions , how did program function or is currently functioning, how much time is needed to move from the inputs to the outcomes and, hence, when impact evaluations can be implemented. It also helps identify what to monitor or measure in order to potentially explain why a program ultimately works or not. In practice, the ToC is a "logical map" of the structure and the expected functioning of the programme or the single project.

The diagrams used to represent a ToC or a programme theory (usually referred to as logic models) can be drawn in different ways:

- as a series of boxes (inputs -> processes -> outputs -> outcomes- > impacts) showing a series of outcomes leading up to the final impacts ('pipeline model');

- in a table; as a series of results, with activities occurring alongside them rather than just at the start.

Software aimed at supporting the visualization of a logic model is also available online and can be very helpful in exercises of this kind.

How to develop a Theory of Change?

The ToC should ideally be developed during the planning stage of a new programme. However, it can also be developed in subsequent times during implementation or at the end of the programme, if there is a need to operate changes in the programme. In the context of the Romanian 2021-2027 programming, the development of the ToC of the ESIF programmes should support the planning of the evaluations and, hence, we can assume that programmes will already be defined at that time.

The definition of the ToC requires a wide range of activities which are often constrained by time or resources available. The main activities are:

- Desk analysis of the programming documents and preparatory studies;

- Review of the literature concerning economic and social theories and empirical evidence related to the funded interventions and relevant past evaluations;

- Interviews or workshops with the programs and the managers of the intervention; and

- Collection of information by the main public and private beneficiaries (e.g., main municipalities, universities, research centers, key companies) and social partners with interests in the interventions.

The quality of the ToC and its future use could be affected if some of the above activities are not implemented; for instance, not involving the main stakeholders in the definition of the ToC could create misunderstandings when they are invited to participate in the evaluation or discuss its findings. It is also possible that a reflection of the elements of the ToC, such as the expected outcomes or the assumptions, has been developed during the design of the programme and now can be "reused".

Developing a ToC for each 2021-2027 Romanian programme would create a "conceptual" map of the programme or interventions, creating an important basis for evaluation planning and programme assessment. The effort to produce ToC of the programmes in a unitary manner is important; to this end, the following simple and straightforward method can be used:

- Reconstructing the intervention logic of the programme based on the official documents by first focusing on the direct relations between instruments and results (input →output →short-term outcomes →long-term outcomes →impacts)[8] and then examining the possible synergies between different results;

1. Identifying the main assumptions by tying together the relationships between different building blocks of the ToC through diverse analytical instruments: literature review and interviews to sectoral experts, meetings with the officials responsible for the programming, meetings with social partners and key beneficiaries;

2. Visualizing and describing ToCs with similar approaches in order to make ToCs comparable. This also implies the definition of similar levels of detail and aggregation for all the OPs.

3. Validating the ToC in a focus group with key representatives from the different groups of stakeholders.

2.4. THE MAIN EVALUATION METHODS

Evaluation research uses many methodologies and techniques developed in social sciences. These methods are not listed extensively in this guide. However, the following features of evaluation research affect the selection and the implementation of methods and techniques:

- **Causality.** Evaluations are interested in studying the causal relations between variables – interventions and their effects. To this end, they largely use statistical or econometric methods when data allows it. Qualitative methods can also be adopted to study causality, as they are capable of describing changes in the behaviors and in mutual relations of individuals, enterprises, or organizations.

- **Comparability and measurability.** Evaluations aim to judge the quality of the examined interventions and, consequently, must measure and compare effects of the same intervention in different contexts or effects of different interventions, targeting the same objective. This implies that the selected methods, quantitative or qualitative, must provide a scale of comparable results and a balanced assessment of the evaluation criteria.

---

8   In reconstructing the intervention logic, using the same structure of the programme can help. Programmes must define the allocation of resources per type of intervention, the specific objectives and the priority or policy objective. In addition, output and results are defined in the indicators and these have their own targets.

- *Types of evaluation.* As mentioned in the introduction, there are several types of evaluations that can be carried out. Also, it is important to note that over time, methodologies have been refined and adapted by type of evaluation. Some methods are mostly used in impact evaluations, while other methods are preferred in process or implementation evaluations.

• **Mixed methods.** Evaluations often address more than one question or use different indicators to respond to a complex question. In these cases, mixed methods, which combine quantitative and qualitative methods can be very useful. These approaches, however, must be carefully designed to ensure that each method targets the right aspect of the problem and overlapping, which would otherwise lead to wasting time and resources, is avoided.

• **Data availability.** Available data counts often for the main constraint in the selection of the methodology. Some statistical techniques are extremely data demanding and the time series or extensive statistics cannot be collected for a single evaluation. On the other hand, qualitative methods generally require detailed field analyses and interactions with many actors to provide robust evidence; these activities are often costly and time consuming.

Even though the objective of this guide is not to provide a comprehensive list of evaluation methods, the following table summarizes the main methodological approaches. Each approach includes an integrated package of options (methods or processes) that can be used in designing a methodology. The most used approaches are:

TABLE 2 Methodological approaches for the main types of evaluations

Type of evaluations	Frequent methodological approaches
Implementation evaluation	- Procedural analyses (procedural journey tracing, IPO Input-Process-Output, data flow diagrams)
	- Market analyses (Surveys on beneficiaries and users, statistical analyses of program outputs and contextual conditions)
Impact evaluations	- Theory based (Theory of change, process tracing, contribution analysis, impact pathways, Realist evaluation)
	- Counterfactual – experimental (Random Control Trials (RCTs), Quasi experiments, Natural experiments)
	- Cost-Benefit Analysis (CBA) on final results ('value for money')
	- Statistical (Statistical modelling, Longitudinal studies, econometrics)
	- Case studies (Ethnography, QCA, within-case analysis, network analysis)
	- Synthesis studies (Meta-analysis, narrative synthesis, realist-based synthesis)

Source: Developed by Authors
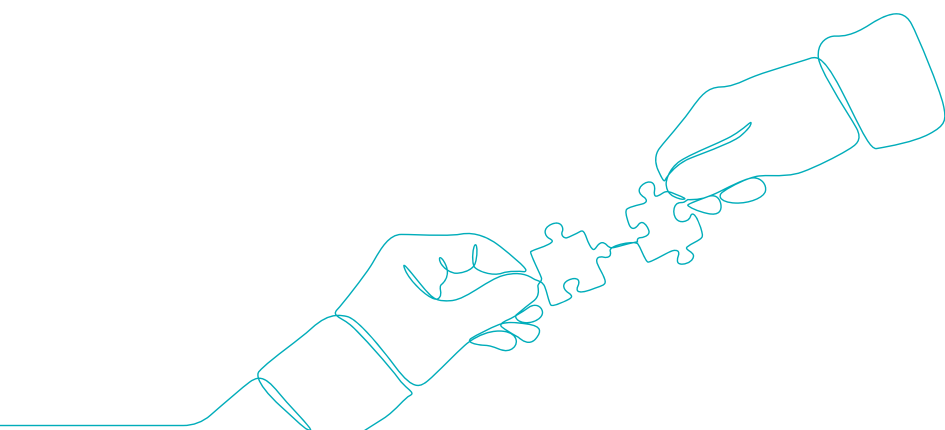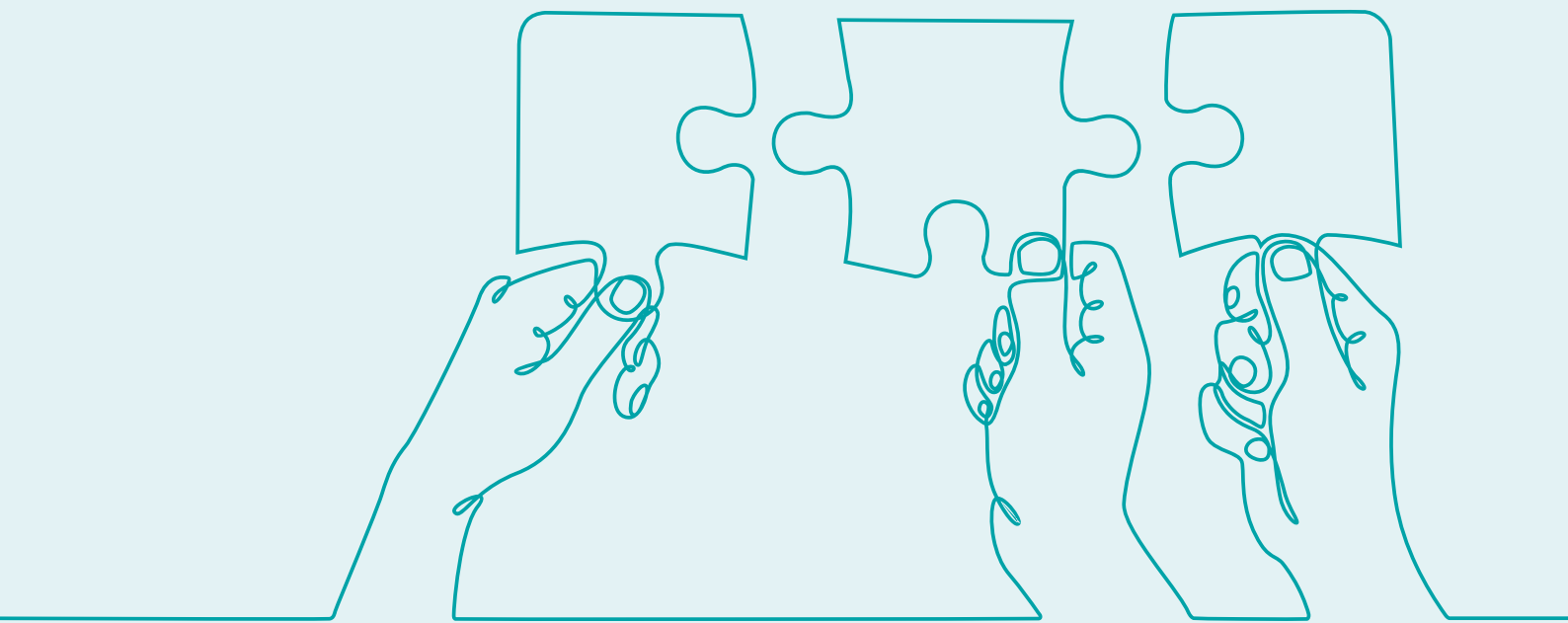
CONTENTS OF THE EVALUATION PLAN

<div align="right">

3

</div>

As shown above, the preparation of an evaluation plan is justified by both regulatory requirements and operational purposes as well as supporting a quality and effective evaluation process. While guidance is available from the EC[9] on the type of information to include in the Evaluation Plan and how it can potentially be structured, the plan must be defined with consideration of the contextual aspects, including the institutional set-up of each country.

An indicative structure is proposed below starting with the EC proposal and past experience in preparing for the 2014-2020 evaluation plans in Romania. The different sections of the Evaluation Plan can be filled in based on the information that will be collected and the arrangements that will be designed in line with the provisions of the current Methodological Document (see chapters 4 and 5 for more details). In terms of format, the Template in Annex 1 should be used.

---

9   Commission Staff Working Document on Performance, monitoring and evaluation of the European Regional Development Fund, the Cohesion Fund and the Just Transition Fund in 2021-2027.
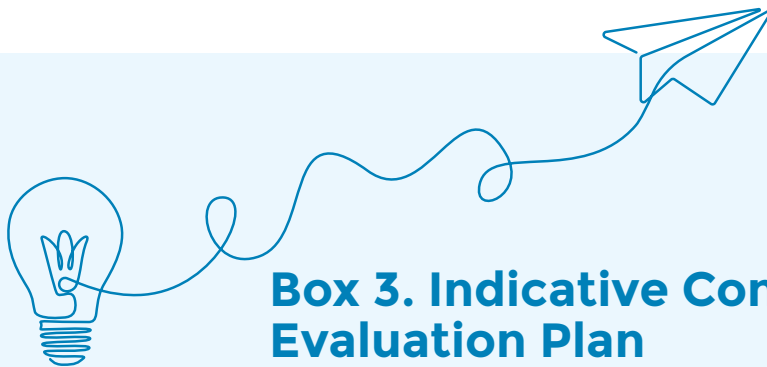
Box 3. Indicative Content of the Evaluation Plan

1. INTRODUCTION

Sets out the Plan's main objectives and describes the key contextual aspects (in terms of scope of OP, institutional architecture etc.). In this section, you must also describe the following:

Coverage and rationale. Explain which OPs and funds are covered by the plan and why. Details on the scope of the OP and key interventions should also be included (ideally these will be supported by the IL or ToC of the OP).

Key processes and roles of different actors in the OP should be also described, to understand areas of interest and potential bottlenecks.

Links with other Plans and Strategies and cross-cutting issues. This section of the Plan will include details on the links with other OPs and relevant national and EU level Strategies. Transversal themes across OPs, common aspects and key interventions that need coordination should be also identified.

Lessons learned and best practices. Includes key takeaways from past 2014-2020 evaluations and studies on the areas covered by the OP. The lessons learned will be used by evaluators and responsible authorities to identify key aspects when improving the evaluation process. Previous evaluations will also affect the selection of evaluations to be carried out, the prioritization of evaluation questions and the definition of evaluation methods and instruments.
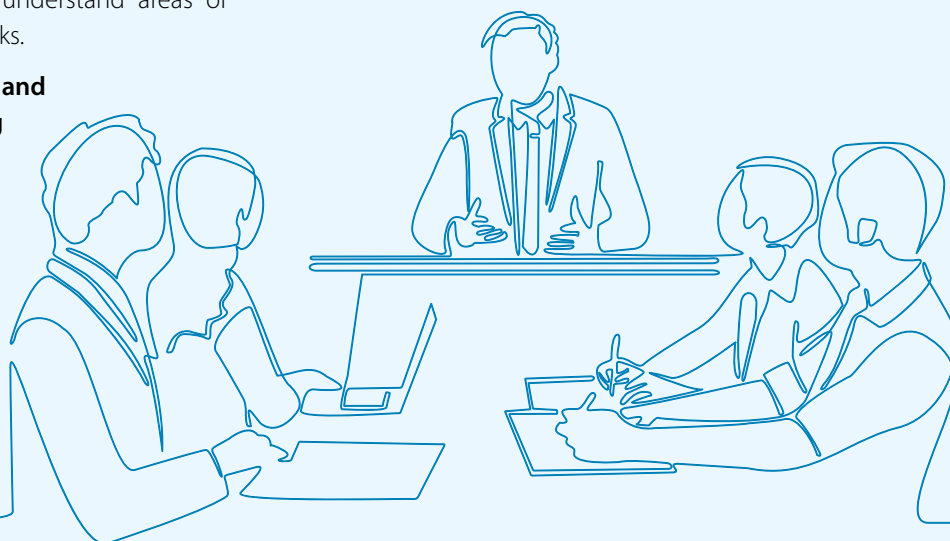
2. EVALUATION FRAMEWORK

This chapter should provide details on the evaluation structures that are set up at PA/OP level for planning, implementing, and managing evaluation and their relationship with coordinating bodies. The structures should also provide a clear picture of the evaluation process led by the Evaluation Central Unit (ECU)/MA and the actors involved throughout the different stages of the process (including capacity related aspects). It should also detail how the evaluations will be sourced and how the results of evaluations will be communicated to the key stakeholders to optimize usage of the evaluations findings. Aspects related to overall budget should also be presented.

The chapter should detail on the following:

a. Role of the MA evaluation unit and its interactions with the key stakeholders

The section will provide detail on the responsibilities of the MA evaluation unit on designing and delivering the evaluation plan and coordinating, monitoring, and promoting the quality of evaluation activities throughout the whole evaluation cycle.

It should also provide information on the decision-making process at MA level (i.e., institutional architecture) and relationship with coordinating bodies (Evaluation Central Unit (ECU) at the Ministry of European Investments and Projects (MEIP) level, line ministries, regional decision-makers, etc.)

b. Involvement of other stakeholders in the evaluation process

This section covers both the arrangements for ensuring the quality of evaluations and those looking towards a greater visibility and awareness of evaluation at the level of the OP monitoring committee, and to promote the use of the evaluation findings.

For ensuring the quality of evaluations, the section should have the following details:

- Any planned steering groups (i.e., Evaluation Steering Committee) and the level at which they will function (for each evaluation, at the level of the Specific Objective, for each OP, at regional/sector level) and intended membership.

- Central technical support from the ECU (main characteristics and aims, beneficiaries of the central support);

- Scientific committees, including details on involving the academia. Details should be provided on coverage (if such committees are to be set-up and used for all evaluations or only for more technical ones), aspects related to contracting (whether they will be contracted by the evaluators or contracted by the MA and used as a mechanism to improve overall quality of evaluations at the OP level), etc.

- Any working groups/networks aiming to ensure a coordinated approach to evaluation and/or promote capacity building. Membership, tasks, and responsibilities should be clarified.

- Responsibilities of each actor (MA Evaluation Unit and groups/committees created) should be detailed, as well as the applicable relationship between them.

Involvement in the framework of the monitoring committee should also be detailed, to increase awareness of evaluation and use of evaluation findings. The section should present details on the following aspects:

- How the results of evaluation will be presented to the monitoring committee members.

- Details on planned working groups responsible for analyzing the evaluation findings, prior to their presentation of monitoring committee.

c. Source of evaluation expertise

Details should be specified on how evaluations will be carried out, whether external, internal or a mix of both. Arrangements for ensuring the impartiality of evaluators for the programme implementers should also be detailed.

It is necessary to include any relevant aspects such as the need for specialized expertise or structure evaluations in a certain manner to facilitate public procurement (i.e., the specificity of areas concerned).

d. Quality management strategy for the evaluation process

The section will provide details on the arrangements for supervising the evaluation process in order to improve its quality. The strategy to this end should cover aspects such as:

- Staff allocated to different tasks.

- Procedures for launching evaluations, appraisal, and selection of offers.

- Supervising the evaluation process (communicating with the evaluators, actions in case of delayed progress, etc.).

- Verification of evaluation reports (procedures, quality checklists, updates to the reports in case of non-compliance with the defined criteria, etc.).

e. Capacity building in the field of evaluation

Knowledge and skills of the staff involved in evaluation are key to a successful evaluation process. This section of the Plan should detail the training program for the staff in the MA evaluation unit and how this will be carried out (e.g., part of a larger training program developed by the ECU, or separately at the level of the MA).

Other arrangements, such as common working groups (at PA level) or professional networks should also be detailed. Coordinating bodies and membership, as well as responsibilities and a functioning schedule should be specified as well.

The integration of the activity of the evaluation unit with other departments in the MA is also key to the relevance of evaluation findings, so the Plan should also provide for arrangements to this end.

f. Communicating the results of evaluation

Communication of evaluation results is a key aspect, having a potentially significant impact on the use of evaluation findings and recommendations. Communication needs to be carefully planned and adapted to different audiences. The strategy to this end should provide details on:

- How the findings and recommendations will be followed up

- How the evaluations will be made public and results of evaluations will be communicated to different types of audiences, to facilitate use and receptiveness. The main categories of audiences include the decision-makers; stakeholders such as MAs, line ministries or European Commission; the general public, etc. Arrangements can refer to the communication mechanism defined, including websites planned/created (i.e., aspects on indicative structure and update mechanism), responsibilities and timeline, etc.

The role of the different working groups and committees in the communication of results should also be detailed.

g. Arrangements for monitoring the implementation of recommendations

Arrangements should be defined for monitoring the implementation of recommendations, stemming either from evaluations of the OP or common, transversal evaluations. These should also be preceded by a screening of the recommendations made by a dedicated body (e.g., steering committee) to decide on which recommendations should be implemented and an indicative timeline. The mechanism could be defined at MA level or implemented by the ECU and its functioning details should be presented in this section.

h. Financial aspects

The overall budget for evaluation, training, and other activities (such as data collection on the side of MA or support for the functioning of the different working groups) should also be estimated and detailed in this section.

3. PLANNED EVALUATIONS

The Plan should include a list and timetable of the evaluation to be carried out throughout the programming period (until June 30th, 2029), as well as an explanation for the selection of the themes covered. The Plan should ensure enough balance between the funds (in case the OP is financed by more than one Fund) and should seek correlation among similar themes and propose possible coordination. The possibility for joint evaluations should be explored for improved effectiveness of the evaluation process and increased relevance of findings.

The Plan should be reviewed regularly by the Monitoring Committee, to accommodate new evaluation needs that may emerge during the programming period. On the other hand, ad-hoc evaluations should remain a valid option throughout the period should the need arise.

The evaluations to be included in the Plan will be determined in line with the guidance provided under Chapter 4 of the current Methodology. For each evaluation, the following details should be specified:

- **Theme/subjects covered by the evaluation** (to be established as per chapter 4, point 2a)

- **Subject and rationale.** It includes background, coverage, the main approach (process and/ or impact evaluation) and the main evaluation questions. All these aspects will be further detailed in the Terms of Reference of the evaluation.

- **Methods to be used and data requirements.** These will be determined in line with point 2b), as presented under chapter 4. They should be detailed in the Plan, as to allow the preparation of the ToR for the respective evaluation (including estimation of the resources needed).

Care should be paid to ensure the right balance between providing enough guidance to the future evaluators on the planned approach and methods to be used for carrying out the specific evaluation and, at the same time, leaving the contractors with enough flexibility to apply their expertise to formulate their own proposals on the best methodology.

Methods planned should also be correlated with the data available (feasible to obtain) as to ensure the evaluators can rigorously implement the planned approach.

Practicality of approach (to avoid evaluations being too scientific and with limited applicability) and efficiency in the use of resources (e.g., optimized selection of methods, use of existing databases etc.) should also be targeted in the process.

- **Data availability.** Details will be presented on the arrangements made to ensure that the specific data needed for the evaluations is available on time. Identification of that data to be collected by the evaluators is also recommended, so that the quality of the evaluations is not affected (in terms of data used, rigor in application of the planned methods, etc.).

- **Duration of evaluation and tentative date for launch/contracting.** The timing can be determined based on the institutional agenda (and corresponding needs for information) as well as regulatory requirements and depend on the type of evaluation (process or impact) and the type of procurement foreseen. The size of evaluation will also influence the timing, to streamline and optimize procurement. Smaller evaluations can be planned on an ongoing basis, while larger, more complex evaluations are expected to be subject to a longer tendering process.

Evaluations should also be planned to allow for the best procurement of expertise (e.g., avoid overcrowding or launching evaluations that are too big), as well as to allow adequate management by the MA Evaluation Units/ECU (i.e., in terms of avoiding overworking staff in the Evaluation Units). Sufficient time must be allowed for procurement and building on the lessons learned from the previous implementation period.
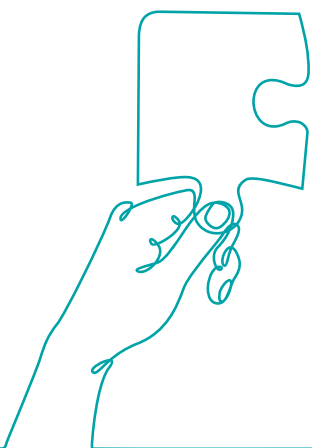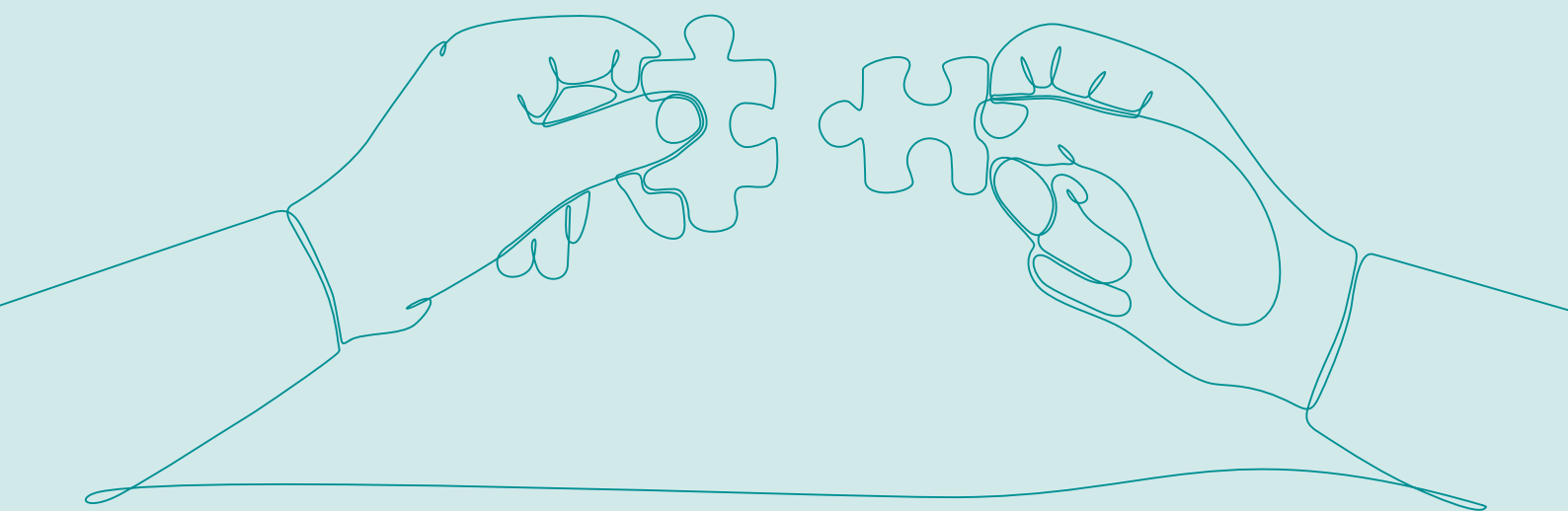
HOW TO IDENTIFY THE EVALUATIONS TO CARRY OUT

4

In this section, a process to identify the evaluations to include in the plan is proposed. It consists of five main sequential steps to implement through several sub-steps and activities (Table 3).

TABLE 3 Five steps process to identify evaluations to include in the plan

Main step	Sub-step	Activity
1) What you want to know	a) Collecting evaluation questions (EQ)	- Mapping the programme (by IL or ToC), identifying the key interventions and the initial evaluation questions of the MA
		- Stating what is already known (desk analysis of past evaluations and literature)
		- Verifying high level strategies and EU and national priorities
		- Involving stakeholders and collecting their questions
	b) Prioritizing and selecting high-level evaluation questions	- Analyzing collected questions (eliminating duplication, grouping, ranking according to the purposes and the regulatory requirements)
2) What evaluations	a) Identifying a set of evaluations to include in the plan	- Associating evaluations to the groups of questions (not too many EQs for each evaluation)
		- Mapping the programme and associating a unit of analysis to each evaluation
		- Verifying evaluations in relation to the regulatory requirements (evaluation criteria, etc.)
		- Defining a horizontal approach (a piece in every evaluation or a 'horizontal' evaluation)
	b) Identifying the appropriate methodological approach for each of the evaluations	- Separating implementation and impact evaluations
		- Identifying main methods for implementation evaluations
		- Identifying main methods for impact evaluations: triangulation of questions, interventions, and methods
		- Verifying data requirements for each evaluation and relative methodological approach
3) When to carry out the evaluation	a) Defining the timing of the planned evaluations	- Verifying institutional agenda and regulation deadlines
		- Paying attention to separate implementation and impact evaluations
		- Verifying the cycle of the intervention to examine
4) Data availability	a) Assessing data availability	- Analyzing available data (main categories and accessibility)
		- Matching data needs and data availability
	b) Identifying possible actions to improve data accessibility	- Facilitating agreements with administrative data owners
		- Promoting a clear and easy way to comply with EU and national rules on data privacy
5) Duration and budget	a) Gathering information on market conditions	- Organizing seminars with market companies and experts to collect information and prepare the market
	b) Estimating duration and budget	- Assessing the needed efforts
		- Identifying the requested skills and expertise
		- Estimating the possible cost

Source: Developed by Authors

The sequence of activities is quite intuitive and similar with other planning processes: identifying the objectives, verifying the best way to reach them, checking the feasibility of the plan, etc. However, in this case, the objective of the plan is "knowledge", which is immaterial, difficult to define and subject to different interpretations. For this reason, the planning process must pay particular attention to what already is known and the opinions of the different actors.

The planning process is defined according to the needs of the MAs, as commissioners of the evaluations, who do not necessarily have expertise as evaluators, but still need to comply with EU rules and produce good quality evaluation plans. As per the EC provisions, the evaluation plan should identify the most suitable methodological approach for the planned evaluations; furthermore, the evaluators will design detailed methodologies in their subsequent proposals based on the initial approach

defined in the evaluation plan. Also, according to the CPR the evaluation plan needs to provide an adequate coverage of the main evaluation criteria (effectiveness, efficiency, relevance, coherence, and EU added value).

At the end of this process and similarly to other planning activities, the evaluation plan must carefully check the feasibility of the identified evaluations to ensure their plan and to successfully implement them. This includes, but is not limited to, checking data availability, estimating time duration and costs, checking if adequate resources are available, etc.

4.1. WHAT YOU WANT TO KNOW

To identify the evaluations that will be carried out, the first step is to determine "what you want to know".

To this aim, MAs must be defined by a set of evaluation questions, as objectives of the future evaluations and based on when the evaluations will be decided. Actionable evaluation questions will be[10]:

- **Evaluative.** They elicit an appraisal of the programme or interventions, according to clear criteria (such as effectiveness, efficiency, etc.) and do not demand a simple description of the policy or the socio-economic context;

- **Pertinent.** They are related to the program's substance and evaluation users' informational needs;

- **Reasonable.** They should examine if a program can realistically achieve the set goals or have an influence on their achievement;

- **Specific.** They must clearly identify what can and has to be investigated in the evaluation;

- **Answerable.** The questions must take into consideration the informative constraints and the data that can be collected, analyzed, and interpreted; and

- **Complete.** The evaluation questions must comprehensively address the purposes of the evaluation and knowledge needs of the users.

4.1.1. Collecting evaluation questions

This section describes the main activities that need to be performed to gather evaluation questions (refer above to Table 3 sub-step 1a) and organize them. These activities include a careful reading of the programme to highlight its main strategic elements, and the examination of the EU and national policy documents. The scope of these activities is to define a first set of evaluation questions. Other activities include consultations with the main programme stakeholders to take onboard their points of view and collect a second set of questions.

4.1.1.1. Mapping the programme and identifying the key interventions

The first activity of the MA and the team responsible for the evaluation is to make a synthetic view of the rationale of the programme, in order to map the planned interventions.

Through the ToC of the programme, it is possible to examine the interrelations between the different interventions and identify the key components of the programme. The importance of individual interventions or coherent sets of interventions will provide a first indication on the evaluation questions and the evaluations that should be conducted. The importance of the interventions derives from a range of characteristics, for instance: some interventions absorb the highest amounts of resources; other interventions work in synergy and are mutually needed to achieve results; there may be interventions that have a key strategic value because of their innovative character, or due to their pivotal role in the programme' strategy. If a ToC has not been yet developed, a simpler reconstruction of the intervention logic (IL) will be useful to interpret the programme.

ToC and IL offer a **visual and logical "map"** of all the interventions of the programme, highlighting where resources are concentrated, the clusters of interventions linked by mutual synergies and common objectives, the sequences of outputs, short-term outcomes, long-term outcomes and impacts to verify. All these elements help to identify the most important components of the programmes and define the evaluation questions.

The Theory of Change (ToC), in particular, identifies the **main assumptions or mechanisms,** which allow the cause-effect chains of the programme to work. These assumptions are hypothetical and need to be verified;

---

10   See Lori Wingate and Daniela Schroeter, *Evaluation Questions Checklist for Program Evaluation*, Western Michigan University – 2016.

this will serve as another important source of evaluation questions. Verifying the ToC of a programme, or part of the ToC related to specific pieces of the programme, is the preferred way to understand why and how some outcomes have been achieved. For very complex or innovative interventions and when the expertise within the administration is not very developed, some external experts (from university or research centres or sector experts) may support this exercise trough a workshop or some interviews.

At the end of the mapping exercise, the following outputs should be available:

1. The identification of the most important interventions (or groups of interventions) and their interactions (operational or strategic synergies, complementarity, temporal sequence of implementation, etc);

2. A better understanding of the operational and strategic issues that need to be tackled in order to reach the medium and long-term objectives of the programme;

3. An initial set of evaluation questions deriving from the analysis of the previous points

4.1.1.2. Examining what is already known

When collecting evaluation questions, there is a need to focus on policy aspects still unknown, to fill the knowledge gaps on policy functioning, rather than repeat evaluations recently carried out.

The evaluation team needs to analyze recent evaluations and academic studies on the topics related to the programme. As part of this process, the main lessons learned must be extracted and the reliability of evidences and methods to identify them should be verified. The results of this analysis eliminate duplication of the evaluation plan with recent studies and will adequately inform the preparation of the next evaluation.

This examination must produce the understanding of the existing knowledge on the themes related to the programme and the main gaps to fill in this knowledge (for instance, the absence of impact evaluations for a certain type of interventions and/or the ignorance on their effectiveness and how they work). This will lead to an evaluation plan that can invest in where there is most need.

4.1.1.3. Verifying high level strategies and EU and national priorities

This activity aims to examine the programme and the main policy documents, and to identify overarching strategies or general policy objectives embodied in the programme. These strategic elements have to be considered in the definition of evaluation questions, in order to verify to what extent the programme contributes to national and EU priorities.

The evaluation team needs to take into consideration the national level or EU level strategies for the 2021-2027 programming period. These strategies may try to provide, for instance, a response to climate change and the decrease of CO_2 emissions, look at compliance with the priorities defined in the European Pillar of Social Rights or aim to regulate management and institutional arrangements during the current programming period. Regardless of their scope, only the strategic themes directly correlated with the programme need to be considered; themes that are too far from the programme action would lead to inconsistent and not pertinent evaluation questions.

This activity will provide a set of evaluation questions related to the main strategic objectives of the programme intends to contribute. The set of questions focus on a few key questions linked to the most important strategic priorities.

4.1.1.4. Involving stakeholders and collecting their questions

Involving the stakeholders in defining the evaluation questions is a key step in drawing up the evaluation plan in line with the needs of the different programme stakeholders and for promoting the future use of the evaluation results. Stakeholders include the following parties: policy makers, implementers, beneficiaries, social partners and other organizations or citizens interested in the effects of the interventions; they offer different perspectives for interpreting the examined intervention; also, there may be cases where their behaviour or actions affect implementation and outcomes of the interventions supported through the programme.

The involvement of stakeholders is further underlined in the EU CPR regulation 1060/2021[11] and requires the participation of social partners and other stakeholders in the Monitoring Committee (art.39). The "Delegated Regulation on the European code of conduct on partnership in the framework of the European Structural and Investment Funds code of conduct" (n.240/2014) dedicated to the functioning of the partnership in ESIF clearly states that MAs should involve partners in the evaluation of the programmes (Art 16). The article specifies that this involvement can occur in the monitoring groups or in specific working groups and the consultation of the partners should also cover *"the reports summarizing the findings of the evaluation".*

A powerful argument for involving relevant stakeholders in the definition of the evaluation questions is that they often know better than the evaluator how the programme works. Use of the information provided by the stakeholders will guide the evaluator or the evaluation planner a more detailed and focused perspective on the most crucial elements of the programme.

While involving the stakeholders in defining the evaluation questions counts for a crucial step in the overall process, the guidance below on how to do this has not a directing character, primarily because of the different institutional arrangements established for the various OPs, but also because of the variety of stakeholders involved. This will provide only an indicative perspective on the approach to be followed. Thus, for increased relevance of this step, it is important however to adapt the approach to the stakeholders' involvement in defining the evaluation questions to the needs and specificity of each OP.

In order to involve the stakeholders, it is first necessary to identify who they are. Relevant stakeholders change based on the chosen intervention to evaluate; for instance, stakeholders in the labor market policy differ from stakeholders in the renewable energy sectors, even if some of them might coincide - for example, – entrepreneurial associations or trade unions which represent general interests. A second element to consider is the role of the different stakeholders; in particular, it is important to identify the stakeholders, who can significantly influence the implementation and the results of the programme.

In general, ESIF programmes have the following main types of stakeholders which, however, may change according to policy sectors covered by the programme and provide different types of knowledge (see Table 4).

---

11   The point 14 of its recital states. *"The principle of partnership is a key feature in the implementation of the Funds, building on the multi-level governance approach and ensuring the involvement of regional, local, urban and other public authorities, civil society, economic and social partners and, where appropriate, research organizations and universities. In order to provide continuity in the organization of partnership, the European code of conduct on partnership for Partnership Agreements and programmes supported by the European Structural and Investment Funds established by the Commission Delegated Regulation (EU) No 240/2014(8) (the 'European code of conduct on partnership') should continue to apply to the Funds.*

TABLE 4 Type of knowledge to be provided by each category of stakeholders

Stakeholder	Details/Sector adaptation	Type of knowledge
Policy makers (political representatives and high-level officials)	Individuals at the top of the decision-making process related to the programme	Main objectives of the programme and expectations from its implementation
The officials responsible for the implementation of the programme in the MA, IB or other authorities	They change according to the sectors and the interventions included in the programme	Operational design of the interventions and implementation issues
Social partners (trade unions and entrepreneurial associations), NGOs or civil associations	In part change according to the sectors and the intervention, but some of them have a general role.	General interests and information (needs, capacities, trends) on important sectors of the society
Other public administrations affected by the programme (municipalities, universities, research centres, sectoral agencies)	Change according to the involved sectors, but some of them (for instance, large municipalities and universities) have a general role (organisational, political cultural, etc.) in the territory	Interest and knowledge on specific territories and policy sectors (e.g., urban and local initiatives, RTDI, energy use)
Public and private beneficiaries (e.g., enterprises, local authorities, citizens of specific territories, etc.)	They change according to the sectors and the interventions included in the programme	Programme implementation, needs to satisfy, expected changes in the behaviours and programme results

Source: Developed by Authors

Some groups of stakeholders may be very large (e.g., beneficiaries) and only a sample of them can be considered; or, some stakeholders are more influential than others on the outcome of the programme and should be consulted first.

An ESIF programme often includes several different policy areas and, consequently, involves several different types of stakeholders; given the pressing time constraints and resource limitations, it is important to select the most important stakeholders and strategize the most efficient way to elicit evaluation questions. In practice, formulating a consultation strategy is necessary.

Various tools to interact with the stakeholders can be adopted, subject to the characteristics of each stakeholders' category. For instance, the following could be used (the list is not exhaustive):

- **Individual or collective interviews** can match the habits of policy makers and officials responsible for the implementation of the programme;

- **A meeting or individual questionnaires** can be used to consult social partners or other influential administrations.

- **Survey or focus groups on samples of beneficiaries** are generally preferable for these numerous stakeholders.

The IL or the ToC can provide a useful guide in identifying the main topics to submit to the stakeholders in the meetings or to include in the questionnaires in order to encourage them to formulate their evaluation questions.

This activity will produce a collection of an additional set of evaluation questions, which will represent different points of views and different interests of the stakeholders in relation to the programme.

The activity will also bring stakeholder participation to the evaluation plan and their "ownership" on the evaluation process. This result will be important during the implementation of the planned evaluations and in the use of their findings. Active participation of the stakeholders in the planning of evaluations is often determinant in the utility of the evaluation results.

4.1.2. Prioritizing and selecting high-level evaluation questions

The numerous questions collected by many different subjects must be arranged and rationalized to make their use possible and effective. This section presents simple and intuitive methods to fall in line with the organization's evaluation questions.

4.1.2.1. Analysis and rationalization of the collected questions

The application of the methods above resulted in a significant number of evaluation questions (detailed in Section 4.1.1), in particular for: (a) the desk research on the structure and the rationale of the programme; (b) the analysis of the overarching EU and national strategies; (c) the consultation of the stakeholders.

The second step of the process is selecting and prioritizing questions; to this aim several screening activities are needed:

1. **Eliminating duplications,** when evaluation questions are redundant and ask the same thing.

2. **Verifying the evaluative nature of the questions.** In case the characteristics of the question do not meet the object of evaluation, they should be assigned to audit or monitoring.

3. **Clustering questions** according to certain important categories and verify if some questions are more pertinent than others. The most important questions can be called **key evaluation questions** (see Box 4); these will provide a backbone of the evaluation plan, while the other questions may be associated or neglected depending on the type of evaluation. Questions can be classified, and their importance can be verified depending on:

   a) *the evaluation criteria* (effectiveness, efficiency, etc., and other possible criteria important in the assessment of the programme),

   b) *the type of assessment required* (questions on implementation and questions on impacts),

   c) *the pertinent policy area or programme components* (questions related to specific priority or policy measures).

This exercise allows for an identification of the key evaluation questions in each group, as well as checking the completeness and the coverage of the collected questions, in comparison to the logic of the programme.

4. **Verifying the feasibility of the most complex questions**; a question may be too complex if it covers too many different measures, takes too much time to measure effects, lacks coherence or is not relevant enough in respect of the programme logic. A feasible question entails that information and methods to answer it are available at a reasonable cost. When necessary, the assessment of the complexity and feasibility of the evaluation questions may be carried out with the help of an external expert.

5. **Checking the usability of the results.** The institutional agenda or programme milestones may require that evaluation findings are available at a specific point in time (e.g., the middle term review, the preparation of the new programming, or national deadlines). There are also cases when some expectations may be particularly high at different administrative levels or among the stakeholders (e.g., the functioning of new policy instruments, the capacity to reach specific target groups, etc.). In practice, some evaluation questions may be preferred to others, because there is an environment that favors the use and discussion of the results in that specific area.
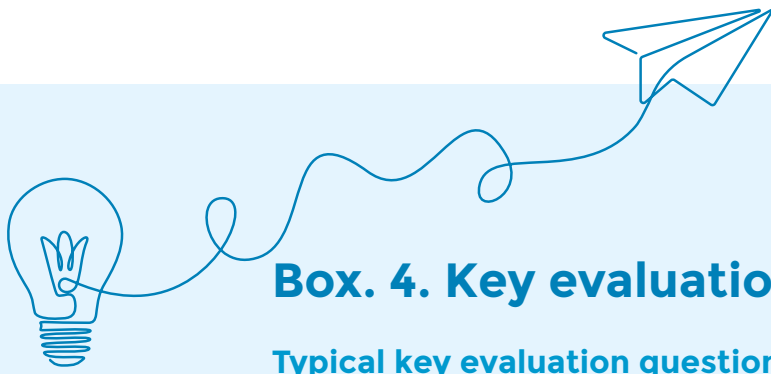
Box. 4. Key evaluation questions

Typical key evaluation questions for the main types of evaluation

Process or implementation evaluation

- How is the program being implemented?
- Is the implementation efficient and timely? if not why?
- How appropriate are the processes compared with basic quality standards? What needs to be changed?
- Are participants being reached as intended?
- How satisfied are program clients? Which clients?
- What has been done in an innovative way?

Outcome or impact evaluation

- How well did the program work?
- Did the program produce or contribute to the intended outcomes in the short, medium, and long term?
- For whom, in what ways and in what circumstances? What unintended outcomes (positive and negative) were produced?
- To what extent can changes be attributed to the program?
- What were the particular features of the program and context that made a difference?
- What was the influence of other factors?

Efficiency assessment (cost-effectiveness analysis and cost-benefit analysis)

- What has been the ratio of costs to benefits?
- What is the most cost-effective option?
- Has the intervention been cost-effective compared to alternatives?
- Is the program the best use of resources?

Source: New South Wales Government, Department of Premier and Cabinet Evaluation Toolkit, https://www.dpc.nsw.gov.au/tools-and-resources/evaluation-toolkit/.

In the annex A of this section, two tools to support the organization of the evaluation questions are proposed:

- A checklist to assess the quality and the importance of an evaluation question and
- A table to list and compare the questions.

These are just examples and can be adjusted according to the needs of the specific evaluation plan.

In selecting the evaluation questions, the MA should respond to what is necessary to know to better implement the programme and make the cohesion policy more effective, but some additional elements should be taken into consideration

1. What is already known and is not necessary to update; and
2. What the EU Regulations (see art. 44 of the CPR) or other national rules on evaluation require.

This activity will produce several and important outputs and principally **a set of feasible and relevant evaluation questions** ordered by

- General relevance ("key" or secondary evaluation questions);
- Evaluation criteria;
- Type of assessment (mainly implementation or impact evaluation);
- Program policy area covered; and
- Usability in relation to institutional and political agenda, as well as programme implementation.

Other important factors to classify the questions may be defined by the MA according to the context, for instance: giving preference to questions related to new or innovative policy measures, or to questions on programme implementation when this is particularly difficult, or to questions coming from important and influent stakeholders.

Even if not compulsory, it would be wise to **validate the selected evaluation questions** in a meeting with the key stakeholders previously involved before proceeding to the design of the evaluation plan. It would allow the MA to work on a shared view of the main issues to evaluate and to anticipate possible disagreement before approving the evaluation plan. This does not mean that the MA has necessarily to find the consensus of all the actors on the evaluation questions, but he/she can try to better consider different points of view and his/her choices can be clear and understandable.

4.2. WHAT EVALUATIONS

After having identified the evaluation questions, the next step involves the identification of the evaluations to carry out in the 2021-2027 period, by coupling the questions, defining the unit of analysis, and verifying compliance with the regulatory requirements. Furthermore, once evaluations are defined, the most suitable methodological approach for each evaluation can be designed.

4.2.1. Identifying a set of evaluations to include in the plan

4.2.1.1. Associating evaluations with the groups of questions

In order for the evaluation to develop a coherent research project and produce reliable results, it has to answer a limited number of key EQs. Too many evaluation questions tend to diminish the effort of the evaluator and make the methodology often too complex and confusing. There is no perfect number of evaluation questions by evaluation, but around 5-7 key EQs is generally a manageable number.

Thus, this activity must cluster the previously selected key EQs into consistent groups. The criteria to group the EQs may differ according to analytical priorities, but some factors are important to consider:

- *Thematic components of the programme covered* (the field of analysis is better defined and more manageable, if questions focus on the same measures or axes; the same is valid for different, but interconnected measures,).
- *Type of questions* (questions can focus on implementation or impact evaluation; questions of the same type allow the evaluation to focus on a specific analysis).

- *Evaluation criteria to which EQs refer* (if questions refer to efficiency, effectiveness, relevance, or other criteria, in order to design a consistent and compact analysis).

- *Political and strategic priorities* (which can give preference to some EQs and require an evaluation of specific policies or interventions).

There is no fixed pattern to combine these factors, which can be merged according to specific needs and priorities. There should be, however, an emphasis on having homogeneous groups if possible.

Grouping should be done only for the "key" evaluation questions, as secondary questions are generally interrelated with key questions; thus, they can be associated, at a later stage, to the more pertinent key questions.

> This activity will identify the groups of homogenous questions, that will be included in the evaluations planned to take place during the programming period.
>
> In this way, a set of key EQs will be associated with each evaluation. The number of key EQs will be not high in each evaluation and they will be, as much as possible, homogeneous by field of analysis and type of required analysis.

4.2.1.2. The unit of analysis and the scope of each evaluation

This second activity consists of defining the right unit of analysis and the scope of each evaluation. This allows to clarify the components of the programme that will be covered by each evaluation, as well as delineate the evaluation methodology.

The unit of analysis represents the level of the programme subject to the evaluation. In general, operational programmes have four main levels: (1) the entire programme, (2) the priority, (3) the specific objective, (4) the actions. These levels are clearly interrelated, but in the specification of the evaluation one level is prevailing.

The scope of the evaluation is defined based on the coverage of the operational programme: one or more priorities, one or more specific objectives, a category of actions etc. It also takes into account the geographic dimension (for instance, only urban areas of a regional programme or only some regions of a national programme, or the same coverage as the programme).

In many cases, the attribution of the unit of analysis and the scope of an evaluation is intuitive, because the evaluation questions immediately identify the right level and component of the programme to investigate. In other cases, it may be more difficult and it may be needed to analyse the ToC (or IL) to specify the unit of the analysis and the scope of the evaluation. In fact, the ToC is a kind of "map" of the programme logic and functioning and helps to identify the pertinent unit of analysis.

It is also noteworthy that a unit of analysis does not prevent the evaluation to investigate different levels of the programme. For instance, strategic evaluations or macroeconomic impact analyses are frequently conducted on the entire programme, but investigations may also require an appraisal of individual priorities or specific objectives or estimating impacts of some large operations. However, the results of these analyses will be provided at the level of the programme and not at the level of its individual components.

> At the end of this activity, the previous selected evaluations will have a clear unit of analysis and scope. Hence, it will be clear to what level, or component, of the programme they refer to and what geographic area they cover.

4.2.1.3. Verifying evaluations in relation to the regulatory requirements

Article 44 of the CPR (1060/2021) provides that the evaluations carried out by MSs should cover one or more of the main evaluation criteria: effectiveness, efficiency, relevance, coherence, and EU added value (for more details see chapter 1), but does not define any additional compulsory requirements to this end. Nonetheless, for an increased relevance of the evaluations, a balanced approach is important.

The definition of the evaluation plan must take into consideration these criteria. The EC does not provide precise indications on how to comply with this requirement, but it clearly represents a general approach

to the evaluation and not an automatic and generalized instruction. There is no obligation that each evaluation covers all criteria in relation to all the components of the OP, as this would imply very extensive evaluations that would be of little use. However, it is important that the planned evaluations overall provide a significant judgment on the quality of the OP factoring in these criteria.

It is necessary to check to what extent each evaluation approaches the different criteria and for what components of the OP. As previously explained, evaluation questions are generally related to the basic evaluation criteria

and, consequently, the evaluation questions of each evaluation provide the logic relationship between the evaluation and the criteria. This appraisal can follow some guiding principles:

- *Proportionality.* It is important to verify that the most financially important or strategic interventions of the OP are evaluated according to the most pertinent criteria (e.g., the intervention of the OP absorbing the bulk of the OP's resources has to be analysed in more detail; if this is deemed to be efficient enough and to clearly respond to needs of the territory, then the evaluation can focus on the effectiveness and coherence of interventions. On the other hand, evaluations of efficiency may be dedicated to interventions that encountered problems in the implementation or have very high costs per unit of output).

- *Timing.* Planning must ensure that the criteria are examined at the right time, subject to the progress in implementation; this will support the effective implementation of the OP and will help respond to informational needs regarding the new programming period or new reprogramming. For example, efficiency may be assessed in evaluations during the implementation period, whereas there is a waiting period until the right conditions are met in order to assess effectiveness. When assessing efficiency, one

can understand if operations are reaching the right targets, their cost is reasonable in comparison to the outputs, and if their implementation is following a planned timetable and are being allowed to absorb EU resources.

Furthermore, the main characteristics of the evaluation are generally connected to some specific criteria, for instance:

- *Implementation evaluations* are mainly focused on efficiency (spending speed, accessibility of the target population to the interventions, cost per output, etc,).

- *Strategic evaluations* aiming to analyse the role of OP in the national context frequently focus on coherence, relevance and EU added value criteria.

- *Impact evaluations* in general are more inherently focused on effectiveness.

It is noteworthy that the CPR also suggests the adoption of inclusiveness, non-discriminatory and visibility criteria. The first two criteria area tend to be especially relevant to social policies, while the latter may be useful to every type of evaluations and, can be assessed during implementation.

When verifying the compliance of the evaluations with the regulatory requirements, the above guidance does not need to be applied "mechanically" but adapted and logically correlated to the evaluation plan.

> The verification of the evaluation criteria used in the planned evaluations will indicate if and how the plan has to be adjusted to better respond to the regulation requirements. For instance, the lack of one of the basic criteria may be justifiable, if past evaluations or other evaluation plans analyse that criterion in-depth on similar interventions. On the contrary, if there are criteria that are lacking and this is not justified by past experience or other reasons, a specific evaluation or the addition of some specific questions to some of the planned evaluations will need to be designed. Such adjustments will ensure the compliance of the evaluation plan to the regulation requirements.

4.2.1.4. Defining a horizontal approach (a piece in every evaluation or a 'horizontal' evaluation)

When identifying evaluations that will be included in the evaluation plan, identifying "horizontal" evaluations is essential, i.e. evaluations aiming to assess the impact of horizontal principles (e.g., environmental sustainability and CO_2 reduction, gender equality, social and political, no discrimination) or effects on territories greater than that covered by the OP (e.g., competitive enhancements at the national level or spill-overs and transregional effects of RTDI operations).

In this case, it is necessary to select one approach between two possible ones:

- A single "horizontal" evaluation, specifically devoted to analyse the horizontal topic in all the interventions of the OP or in all the pertinent OPs. This approach needs a dedicated and probably quite a sizable evaluation.

- An evaluation combining contributions (for the specific horizontal theme) deriving from the evaluations carried out in the context of single OPs.

This approach requires a good coordination among the MAs evaluation units/ the bodies responsible for evaluation of the different OPs in order to have the contribution from different OPs at the same time, according to comparable methodologies.

These two approaches are not necessarily mutually exclusive and they may be combined by having a horizontal evaluation which receives contributions from different programme evaluations and also produces original analyses and evidence. However, the operational and coordination effort of this combined approach is probably very extensive.

This activity will allow positioning the evaluation of the horizontal questions in the right way inside the plan, as well as anticipating coordination and operational issues related to the implementation of this kind of evaluations.

4.2.2. Identifying the appropriate methodological approach for each of the evaluations

After deciding on the evaluations to be implemented, it is necessary to identify the most appropriate methodology for each evaluation. As per the applicable EC requirements, the evaluation plan should identify the preliminary methodological approach for the evaluations, to support the preparation of the call for tender, as well as estimate the resources and time needed. By methodological approach we mean a package of integrated and coherent methodological options (methods or processes)[12]. A more detailed and operational design of these methodological options will be delineated during the call for tender or will be expressly requested to the tenderers.

When selecting the methodological approach of an evaluation, it is important to highlight the following fundamental concepts:

Evaluation focuses on the causality between the public policy (e.g., an expenditure, a new rule or institution, a collective decision) and the changes in the socio-economic context; hence, the evaluation methodology has to analyse causalities and relations between social phenomena.

A public policy, or an operational programme, is designed to trigger a complex set of cause-effects relations influencing social behaviours, investments, employment, development patterns, etc. The ToC, or the Intervention Logic, summarizes these causalities and their expected effects; the evaluation methodology must verify to what extent these causal chains are factual and affect society in the desired direction.

It is important to understand that the results of any public policy, and especially the results of a complex operational programme, do not function in a linear and mechanical way, whereby a public expenditure automatically produces the expected results. If this were true, we would not need the evaluation. On the contrary, effects of public policies may be highly indeterminate either because their initial design may be vague at times or because they are influenced by a multitude of external factors such as how policies are implemented and macro-economic or institutional conditions. Consequently, the evaluation methodology should support and distinguish to what extent the policy made the difference, or how the contextual factors determined the social changes. The separation between the influence of the policy and the influence of the contextual factors is a huge challenge for the evaluation methodology since it cannot reproduce the society in a laboratory.

These considerations explain how difficult it is to define a fully successful methodology and why evaluation findings often have limitations and cannot be clear-cut. These considerations are also the foundations of the methodological design, which is the result of the cross-reference of three elements (Figure 4):

- *The evaluation questions,* which determine what effects of the intervention have to be examined (some questions ask "what" results, others "why" these results, etc. and each of these questions needs different types of analyses);

---

12   The most used and important methodological approaches have been presented in section 1.

- *The characteristics of the intervention to evaluate*, which determine the type of causalities to analyze (e.g., local development interventions involve many different actors and activate multiple types of causality, while subsidies for investments involve similar enterprises and trigger a specific effect on the firms' decisions).

- *The available methodological approaches,* which are the "toolbox" where the most appropriate method can be selected to answer the evaluation questions.

FIGURE 4 Key elements for defining the methodological approach
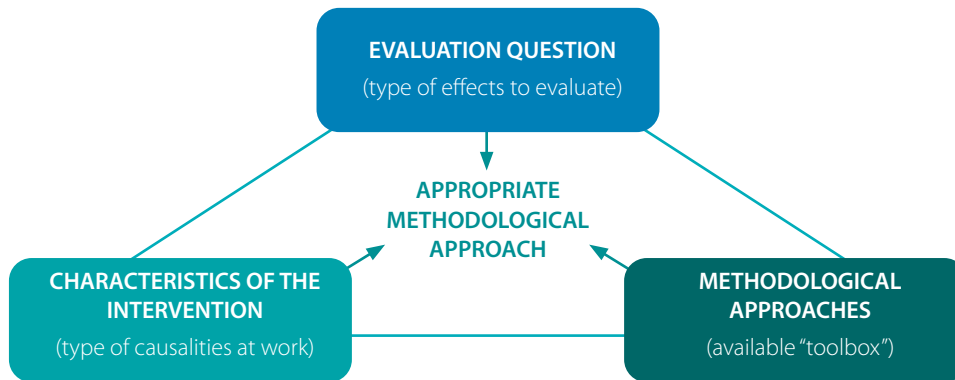
Source: adapted from Stern, E. (2015). Impact Evaluation: A Guide for Commissioners and Managers. Prepared for the Big Lottery Fund, Bond, Comic Relief, and the UK Department for International Development.

It is noteworthy that many evaluations use mixed-method approaches. They integrate different methods in the same study, such as quantitative and qualitative analyses or quasi-experimental counterfactual methods and theory-based methods. This solution may be needed when evaluation questions focus on different aspects of the outcome and consequently the evidence that need to be collected significantly differ. In other cases, results of the interventions are ambiguous. In this case, there is a need for more solid proofs, which – in turn – require the combination of different methodological approaches to achieve a proper assessment. Furthermore, when using the mixed-method approach, it is important to link each method to a specific analysis and avoiding overlapping or useless investigations.

These considerations are general and have theoretical implications, which we do not aim to develop in the framework of the current guide, but they are important and orient the methodological choices. The following paragraphs suggest some key steps that can be followed to define the mostly suitable methodological approach for a specific evaluation.

4.2.2.1. Separating implementation and impact evaluations

The first step would be to separate the evaluations focused on the implementation of the OP or its interventions from the evaluations aiming to assess the final results and the impact of the OP. The separation creates two groups of similar types of evaluation and confines the methodological definition in two homogeneous fields.

The classification of the evaluations in two groups is intuitive, but evaluation questions can also guide the clustering:

- *Implementation, or process, evaluations.* These generally respond to questions such as: the efficiency of the implementation processes, the types of beneficiaries involved and the relevance of interventions to the objectives of the programme/intervention, the importance and the coherence of the programme's expenditure in relation to other national policies, the correctness of the initial strategy, including the context of potentially significant changes in the socio-economic and policy context, etc. These evaluations are usually 'ongoing' as there is no need to wait for the end of the interventions to carry out such evaluations and normally, they provide findings that can be used during the implementation of the programme.

- *Impact evaluations.* These evaluations generally respond to questions such as: what are the final and net effects of the programme, was the programme able to make the difference and did it reach its objectives, how the programme succeeded in

producing its results, whether there are unexpected or undesired effects that have occurred, what is the "value for money" of the impacts, etc. These evaluations are carried out in the final stage or at the end of the interventions, or even some time after the finalization of the programme, so that the effects take place and can therefore be examined. If the effects of some measures are visible only after the closure of the OP13, it is possible to evaluate the impact of similar interventions completed in the previous programming period. For the results from old interventions to be useful for current interventions, analyzed results need to be generalisable. This means that the interventions financed during the two periods must be very similar, if not the same. Moreover, the evaluation methodology must take into consideration the possible influence of changes in the socio-economic context.

It may be possible that an evaluation includes analyses on the implementation as well as on the impacts. In this case, the attribution of the evaluation to one or another group will follow logical considerations, such as the prevalent analysis.

> This activity will produce two groups of planned evaluations (implementation evaluations and impact evaluations), which are homogenous in relation to the type of effects to look for and, consequently, pose similar methodological problems.

4.2.2.2. Identifying main methods for implementation evaluations

The identification of the methodological approach for implementation evaluations should take the following into account:

- Specific aspects to examine (procedures, financial allocations, type of beneficiaries, outputs, etc.)

- Scope and the size of the analysis (territorial scope, sectoral scope, the entire OP or just a component, etc.).

Table 5 below proposes an example of these elements associated to the most frequently used methodology.

TABLE 5 Implementation evaluation: evaluation's object and methodology

Object of the evaluation	Examples of methodological approach	Scope and size of the analysis
Management procedures *(e.g., are interventions in time? Are they efficient? What are the administrative costs and the idle times? etc.)*	- Procedural analysis - Interviews of the management staff - Survey of the selected and rejected applicants - Case studies of specific procedures	If many procedures or interventions have to be examined, a survey on the implementers may allow an easier coverage.
Types of beneficiaries and funded projects *(e.g., Is the programme reaching out its target? Are "creaming" effects at work? Etc.)*	- Survey on the beneficiaries - Desk analysis of the projects - Networking analysis - Case studies	When beneficiaries are many and similar (e.g., enterprises) a survey may be suggested
Strategy of the programme/ intervention *(e.g., is the OP's strategy relevant? Are interventions producing added value in comparison to national interventions? Etc.)*	- Statistical analyses of the programme outputs and related socio-economic variables - Desk research on functioning and objectives of related national policies	When an entire national OP is analyzed, macro-economic aggregate may probably be used in correlation to outputs and outcomes. For sectoral or territorial analyses, more detailed data must be available.

---

13   This may be the case of some infrastructures of interventions which are completed just at the end of the programming period. On this point see also the next section 4.3.1.3 *"Verifying the cycle of the intervention".*

Object of the evaluation	Examples of methodological approach	Scope and size of the analysis
Potential effects of the programme *(e.g., Are outputs and initial results in line with the ToC? Are the expected targets and benefits reached? What is the cost-effectiveness of the intervention?)*	- CBA - Statistical analysis of the (actual or expected) output	CBA can be effectively applied to specific and large projects. A large number of small projects or individual effects would probably require statistical analyses.

Source: Developed by Authors

In the implementation of evaluations, surveys are frequently used, justified by the large number of subjects to contact. It is risky to survey only one category of actors (beneficiaries or implementers) as this results in biased design. Analyzing the behaviors of only one actor and not considering the different points of view – for example, the point of view of both implementers and beneficiaries – shows only a partial picture of the situation. Thus, it is important that the "triangulation" between the different opinions at stake takes place.

Another frequent pitfall in the implementation of evaluations is the use of monitoring data alone. This approach adds very little to the available knowledge and it will be difficult to come with novel elements in judging the progress of the programme. Monitoring information is important for the evaluation, as it allows evaluators to understand what has been done and when. However, it has to be included in a broader, evaluation-specific methodological design. In addition, monitoring is only one source of information, collected by the administration during the implementation. For instance, applications generally contain a lot of data regarding the funded projects and their planned outputs, data on output and result indicators, as well as data on spending, which is collected throughout implementation. The analysis of wider administrative data should be associated with the analysis of monitoring data.
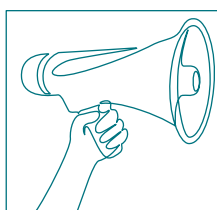
This activity will define an appropriate methodological approach for each planned implementation evaluation.

4.2.2.3. Identifying main methods for impact evaluations

Unlike implementation evaluations which focus on aspects influencing implementation, impact evaluations analyze effects; impacts are significantly "distant" from the implementation, both from a temporal perspective and when looking at the intervention logic. They are distant in temporal terms as there is a need for the analysis to occur after a certain period of time once the intervention ends. They are also distant in logic terms and require that result chains are properly identified (i.e., immediate and intermediate results are chained in the right way) in order to understand *why* something worked or didn't work. In practice, outcomes and impacts are complicated, with some being distant from implementation and/or significantly influenced by external factors.

The selection of the appropriate method will be based on the following:

- Cross-analysis of the evaluation questions
  - Characteristics of the programme/intervention
- Potential methodological approaches

Table 6 below shows the characteristics of the main methodological approaches to impact evaluation. The table highlights the types of causal inference that these methodological approaches best examine, as well as the main strengths and weaknesses of each approach.

TABLE 6 Impact evaluation: main methodological approaches and causal inference

Design approaches	Specific variants	Basis for causal inference	Strength and weak points
Experimental	Random Control Trials (RCTs), Quasi experiments, Natural experiments	**Counterfactuals:** the difference between two otherwise identical cases – the treated and the controlled;	**Strength** – good measurement of the effects **Weakness** – limited generalization of the results
Statistical	Statistical modelling, Longitudinal studies, econometrics	**Regularity:** Correlation between cause and effect or between variables, influence of (usually) isolatable multiple causes on a single effect. Control for 'confounders' affecting both causes and effects.	**Strength** – based on significant relations **Weakness** – a high n. of cases is necessary in statistical analyses. The results of models largely depend on internal assumptions
Theory-based	**Causal process designs:** Theory of change, process tracing, contribution analysis, impact pathways **Causal mechanism designs:** Realist evaluation, congruence analysis	**Generative causation:** Identification and confirmation of causal processes or 'chains'. Supporting factors and mechanisms at work in a specific context.	**Strength** – high explanation capacity of the functioning of the intervention **Weakness** – not precise measurement of the effects
Case-based	**Interpretative:** Naturalistic, grounded theory, ethnography **Structured:** Configurations, Qualitative Comparative Analysis (QCA), within-case analysis, simulations, and network analysis	**Multiple causation:** Comparison across and within cases of combinations of causal factors. Analytic generalization based on theory.	**Strength** – effective with a limited n. of projects or beneficiaries **Weakness** – difficulty with high complexity and many different possible causes
Participatory	**Normative designs:** Participatory or democratic evaluation, empowerment evaluation. **Agency designs:** Learning by doing, policy dialogue, collaborative action research.	**Actor agency:** Validation by participants that their actions and experienced effects are 'caused' by programme adoption, customization, and commitment to a goal	**Strength** – direct observation of the effects on the beneficiaries **Weakness** – potentially biased by subjective views
Synthesis studies	Meta-analysis, narrative synthesis, realist-based synthesis	**Accumulation** and aggregation within a number of perspectives (statistical, theory based, ethnographic.)	**Strength** – findings are quite well generalizable **Weakness** – availability of many comparable studies is needed

Source: adapted from Stern, E. (2015). Impact Evaluation: A Guide for Commissioners and Managers. Prepared for the Big Lottery Fund, Bond, Comic Relief, and the UK Department for International Development.

The causal inference and the strengths and weaknesses of each method influence the way the respective method can be associated with the evaluation questions and the type of intervention selected to use for evaluation. Table 7 below proposes some of these possible associations. These associations are indicative and cannot be applied automatically. Selecting the methodological approach for an evaluation question or intervention always requires a careful examination of the different elements that indicate what type of analyses should be carried out.

The choice of the methodological approach may be affected by the combination of different evaluation questions or by unclear questions; in these cases, it is possible to identify a methodological approach for each key evaluation question. When evaluation questions are unclear and do not allow the selection of the method, it is advisable to review the questions. The selection might be difficult in the absence of the ToC of the programme or the intervention. In this case, if it is not possible to work on the ToC, the selection can be generic, and a more detailed selection could be done at a later stage when the evaluation should be implemented. It is also important to remember that it is the task of the evaluator to define the detailed methodology.

> This activity will identify the suitable methodological approach to each planned impact evaluation.

TABLE 7 Evaluation questions, type of intervention and methodological approach

| Key questions | Related evaluation questions | Intervention | | Suitable designs |
		Examples	Characteristics	
To what extent can a specific (net) impact be attributed to the intervention? (No particular interest in generalization)	- What is the net effect of the intervention? - How much of the impact can be attributed to the intervention? - What would have happened without the intervention?	- Subsidies for enterprises - Large training schemes - Large interventions in education - Infrastructural investments - Urban policies - Social inclusion interventions	- Sufficient numbers (beneficiaries, households, etc.) for statistical analysis - Expected outcomes clearly understood and specifiable - Limited number of cases - Interaction of several potential causes	- Experimental (counterfactual) - Statistical studies - CBA - Hybrids with case-based and participatory designs
Has the intervention made a difference? (Interventions are just one part of a causal package)	- Was the intervention needed to produce the effect? - Would these impacts have occurred anyhow? - What causes are necessary or sufficient for the effect?	- Subsidies for enterprises - Large training or education schemes - Several and similar public works comparable in vary territories - Local development initiatives - Urban policies - Social inclusion - R&D infrastructures	- High number of similar cases (beneficiaries, projects, or participants) - Comparable cases with a common set of potential causes - Several causes that need to be disentangled	- Experimental (counterfactual) - Theory-based evaluation, e.g., contribution analysis - Case-based designs, e.g., QCA

Key questions	Related evaluation questions	Intervention		Suitable designs
		Examples	Characteristics	
How has the intervention made a difference?	- How and why have the impacts come about? - What causal factors have resulted in the observed impacts? - Has the intervention resulted in any unintended impacts? - For whom has the intervention made a difference?	- All types of intervention	- Interventions interact with several other causal factors - A theory that allows identifying supporting factors (proximate, contextual, historical, etc.) is needed	- Theory-based evaluation (especially 'realist' variants and Contribution Analysis) - Participatory approaches
Can this be expected to work elsewhere? (What has worked in one place/time and can work somewhere else or in the future)	- Can this 'pilot' be transferred elsewhere and scaled up? - Is the intervention sustainable? - What generalizable lessons have we learned about impact?	- All types of intervention	- Generic understanding of contexts e.g., typologies of context - Clusters of causal packages - Expected innovation diffusion mechanisms	- Mixed-method approach (Participatory + some Experimental and Theory-based Approaches) - Realist evaluation - Synthesis studies

Source: adjusted from Stern, E. (2015). Impact Evaluation: A Guide for Commissioners and Managers. Prepared for the Big Lottery Fund, Bond, Comic Relief, and the UK Department for International Development.

4.2.2.4. Verifying the methodological approach

A verification of the methodological assignments can help to fine tune and validate the choices. Some key factors should be verified:

- **Provision of the right evidence.** The selected methodology has to provide "actionable" evidence to answer the evaluation questions; e.g., counterfactual approaches provide a measure of the 'net effect' of the intervention, case studies deliver information on the efficiency of the intervention and the socio-economic changes that have occurred in some specific contexts. In order to check the matching the methodological approach and the evaluation questions, there is a need to verify to what extent the potential indicators defined through the methodology are adequate to answer the questions. In case of mixed-method approaches, it also helps to verify if two methods produce the same, or very similar evidence. In such a case, one of the methods can be left out.

- **Feasibility and data availability.** The feasibility of the selected methodological approach is an additional fundamental check. The feasibility can depend on the time requested to use the methodology (e.g., survey takes time, but case studies can take more time), the needed skills (e.g., carrying out counterfactual impact evaluations requires specific knowledge on econometrics and statistics) and, data availability. If the required data is unavailable and cannot be readily produced, the selected methodology is not viable or can encounter serious problems during the implementation (e.g., availability of data for a control group in a counterfactual evaluation). The next sections explain in more detail what type of

data is generally used in ESIF evaluation and its availability; at this point, it is important to underline that data availability is a crucial factor to verify the methodological choices and their reliability.

- **Feasibility and available resources.** The cost of a methodology is another crucial factor where there is a need to control in advance. Some methodological approaches may be too expensive compared to the available resources; for instance, surveys have relatively standard costs and can simply be estimated; other methods can be more complex, but their costs can be estimated by using uncomplicated market analysis or by conducting interviews with practitioners. The guide provides more details in the following sections.

- **Ethical constraints.** Some methodologies can have significant ethical implications which require paying particular attention in their selection. For instance, the use of randomized trial control approaches may lead to ddiscrimination against participants, when the selection is based on the random probability; also, using individual data may conflict data protection rules if not properly managed. Discriminatory practices should be avoided in case studies or in the formulation of the questionnaires, or copyright laws could limit the access to information on enterprises. The gravity of these issues can vary according to the context, but, in general, preparing for them in advance prevents the need for changing methods during the implementation phase.

4.3. WHEN TO CARRY OUT THE EVALUATIONS

The timing of evaluations is another important choice in the definition of the evaluation plan. The following paragraph introduces some basic factors to be considered in setting the timetable of the plan.

It is important to highlight that the time needed for completing an evaluation may be significantly affected by the duration of the public procurement procedures for selecting the evaluators. In the past, these procedures sometime took more than one year, so there is a need to find solutions to accelerate the process in future. However, in the planning of evaluations, sufficient time should be allowed for the procurement procedures, as well as for carrying out the evaluation itself.

4.3.1. Defining the timing of the evaluations

4.3.1.1. Verifying the institutional agenda and regulation deadlines

The first element to consider is the national and EU policy agenda of the OP. If some institutional deadline or event requires the preparation of a specific evaluation, its timing needs be fixed as a milestone in the plan.

With reference to the article 44 on evaluation in the Common Provision Regulation n.1060/2021 some deadlines are defined for the MS:

- For ERDF, CF and ESF, an impact evaluation of each OP must be carried out by June 30, 2029, to take stock of the funded interventions and prepare the new programming period.

- in the case of the Asylum, Migration and Integration Fund (AMIF), the Internal Security Fund (ISF) and the Instrument for Financial Support for Border Management and Visa Policy (BMVI), the evaluation plan has to include a mid-term evaluation to be completed by March 31, 2024.

In article 45, the EC is required to carry out:

- A mid-term evaluation to examine the effectiveness, efficiency, relevance, coherence, and Union added value of each Fund by the end of 2024.

- A retrospective evaluation of each Fund by December 31, 2031.

Hence, MSs are not obliged to carry out a mid-term evaluation, but a coordination of national evaluations with the evaluation carried out by EC may be useful to ensure a wide information base to the European exercise.

Other deadlines may come from the national agenda or given the need to coordinate some key evaluations at the national level (for instance, a mid-term review of the issues related to implementation or the estimation of the initial impacts in some key policy areas).

4.3.1.2. Paying attention to separate implementation and impact evaluations

A second suggestion regarding the difference between implementation and impact evaluations:

a. Implementation evaluations need to be concluded during the implementation and in time to use their results for corrections in the operational procedures or reprogramming (changing the OP).

• Impact evaluations need to be carried out when the effects of the interventions are available and this may occur years after their conclusion (e.g., private investments produce effects in the enterprises after 1-2 years, generally). However, if information on the

impact of some specific interventions is important and urgent in order to implement the OP and these types of interventions were already completed in the previous programming period, it is possible to focus the impact evaluation on those interventions, paying attention to possible changes in the socio-economic context that could differently affect the new interventions14.

Hence, an initial division between implementation and impact evaluation can help to clearly identify the timing of the evaluations.

4.3.1.3. Verifying the cycle of the intervention

Regarding impact evaluations, it is important to highlight that not all the interventions have the same "time to impact". On the contrary, important differences apply regarding the time needed for impacts to take place.

In terms of the time needed from design of the intervention to its impact, it is recognized that training

and active labour market policy can generate impact in around 1- 1,5 year, while infrastructures can take 15 years to cause some impacts. In between these two extremes, we have many other interventions like RTDI or subsidies to private investments (Figure 5).

FIGURE 5 The time needed to produce impacts
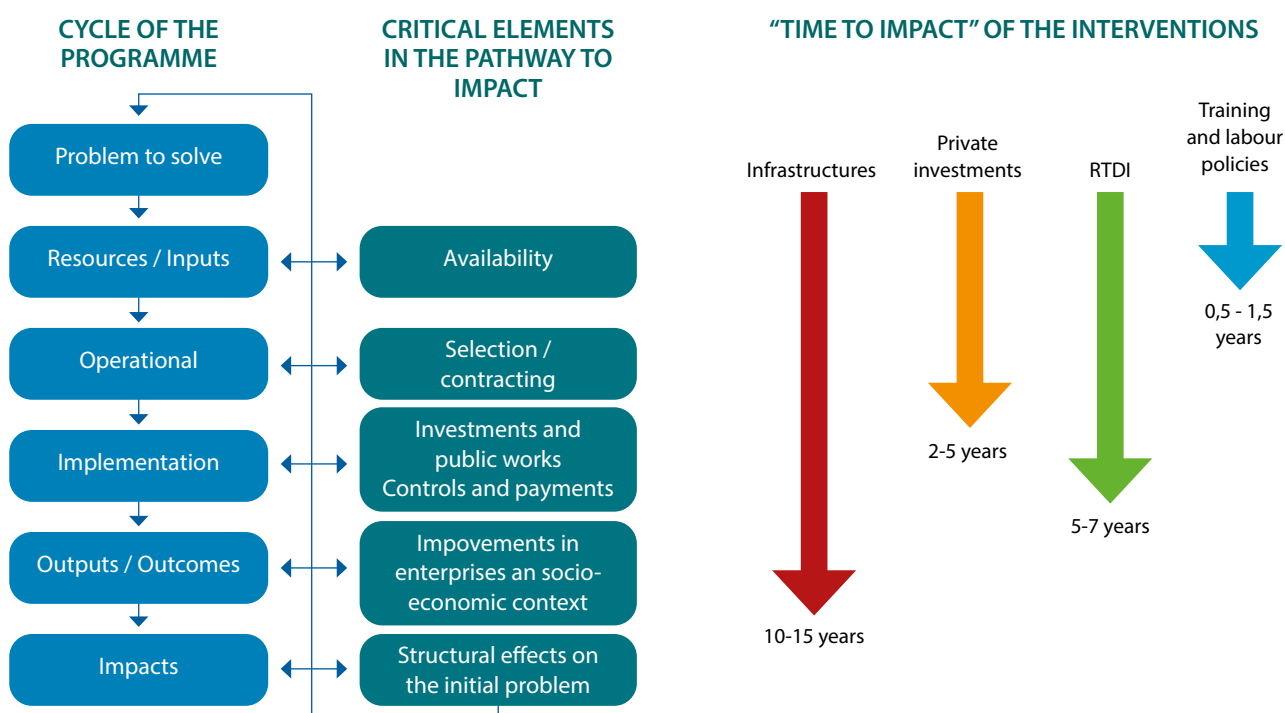
Source: Developed by Authors

---

14   This approach has often been followed in the 2014-2020 period, when the initial impact evaluations sometimes examined 2007-2013 investments.

These indications have an explanatory value; in practice, the right time to analyze the impact of an intervention must be calculated by considering the characteristics of the individual intervention. In this calculation, it is also important to carefully distinguish outputs, outcomes and impacts; they occur in sequence and their timing can vary according to the timing of reactions among the final and indirect beneficiaries (e.g., the impact of a new metro line does not require only its functioning, but also to be fully accessible by potential users; changes in the behaviour of the people living in the area are also needed for impact to occur). In addition, it is noteworthy that some interventions, such as large infrastructure, are selected for financing under the OP, when they are already in an advanced stage of preparation; for instance, large infrastructures are generally selected for financing under EU programmes when projects are completed and permissions are issued; the timing of their impacts have to be calculated from that stage, not from their start.

Considering the above, for each type of interventions included in an evaluation, it is important to follow some basic steps in order to define when that tender for that evaluation should be launched:

1. Appraising the stage, the intervention is, at the moment of its inclusion in the OP (planning, designing, selecting the beneficiary, executing the investment, using the results of the investment, etc.) and the time when it was or will be included in the OP;

2. Estimating the time that the intervention needs to produce outputs, outcomes or impacts (according to the objectives of evaluation to carry out) from the moment it is included in the OP and calculating the relative approximate date for these to occur;

3. Estimating the time needed to select the evaluator and to carry out the evaluation;

4. Subtracting the time needed to select the evaluator and to carry out the evaluation from the date when outputs, outcomes or impacts will be available. Only after this final step, we can determine the date when the evaluation has to be launched.

4.4. DATA AVAILABILITY

High quality evaluations require that high quality data is available to the research team in a timely manner. An evaluation can rely on one or more data sources which can be both quantitative and qualitative.

Evaluation teams can rely on existing data and/or collecting their own information. Existing data, including administrative sources or official statistics, is generally less expensive and more easily accessible than collecting the information. However, it is not always possible to substitute one source with another, since data may be defined, collected and organized differently; in practice, the same topic may have different data sources.

In the following paragraphs, basic guidance on how to check data availability including some indicative activities are mentioned; in the context of this guide, the information provided will not cover all the characteristics of each data source. These characteristics are not always available in public documentation.

4.4.1. Assessing data availability

4.4.1.1. Analyzing available data: main categories and accessibility

The statistical database potentially needed for an evaluation includes information collected from official administrative sources, such as monitoring of the intervention and public registers, and from original sources, where information is collected directly through field surveys (relying on individuals, companies and other organizations). Data sources can be broadly categorized into three types:

1. **primary sources.** These involve those that are collected first-hand by the evaluation team.

   Data can be collected at different units (or levels), including for example individual, household, company, or a geographical area. Data may be collected from both beneficiary and non-beneficiary units, especially in the case of impact evaluations where a counterfactual is needed. Primary data sources often add to other data sources such as monitoring data. This data is often used both in implementation and impact evaluations to add to the information from the monitoring system and to have micro-level information on final beneficiaries or control groups in counterfactual analyses.

2. **secondary administrative sources.** Two main types of secondary administrative sources are relevant to the evaluation:

   a) monitoring data that directly relates to the policy

or intervention being evaluated. This can include financial and physical (operational) data.

b) data from administrative registers (e.g., registers of employment and social security for workers, merit of credit and tax registers for enterprises and individuals, etc.).

The access to administrative data is subject to data protection rules, which requires specific procedures and agreements between the MA and the data owners. This is the reason potential data providers should be contacted early in the process to gauge a variety of factors like the timing and access to the type and quality of data.

3. **secondary statistical sources**, which provides more detailed information at aggregated and microlevels on final beneficiaries (companies or individuals) and similar subjects; in addition, these sources provide fundamental statistics on the context in which they operate. The data derives from the official statistical institutes or from the organizations which produce public and reliable statistics (e.g., national bank, house

of commerce, public or private research institutes).

Data is fundamentally important for investigating the context in which the intervention is implemented, for estimating macro-economic effects and for following the trends of aggregated variables related to the situation of the final beneficiaries (e.g., investments, productivity, transported commodities and passengers, R&D expenditure, enrolled students, etc.). In general, the objective of the analysis is twofold: understanding the extent to which the intervention modifies the context and socio-economic trends, and the extent to which contextual changes affect the outcomes of the intervention. The access to statistical sources is generally free, but this type of data suffers from limitations in geographic and sectoral details. In some cases, when it is possible to access micro-data of the official surveys (Labour Force Survey - LFS, EU Statistics on Income and Living Conditions – EU SILC etc.), this data can be also used for counterfactual impact evaluations on enterprises or individuals.

4.4.1.2. Matching data needs and availability

Based on the evaluation questions discussed above, it is possible to identify six main categories of indicators necessary for the evaluation, constructed based on information available from different sources:

1. **Indicators and variables necessary for describing the intervention.** These derive from the monitoring of the programme and inform how the programme/ intervention is performing and what are its outputs.

2. **Procedural and administrative indicators.** These describe the characteristics and the efficiency of the implementation system. These indicators are necessary to evaluate the management system in the framework of implementation evaluations and are also useful in impact evaluations to analyze to what extent implementation affected final outcomes;

3. **Outcomes and impact indicators**, which are needed to answer the evaluation questions related to impact (generally: what? why? how?).

4. **Qualitative variables and indicators** capture the beneficiary's appraisal of the effectiveness and the functionality of the intervention. These indicators may be fundamental in implementation as well as in impact evaluations. The information on these indicators is collected by asking beneficiaries or stakeholders how they react to the intervention and

how the intervention changed their behaviors in general or in their relations with other actors;

5. **Counterfactual evaluations** indicators look at the variables that are used to define the matching (or similarities) between 'treated' and 'control' units (between those benefiting and not benefiting from an intervention). This information may be collected by surveys or by administrative data sources.

6. **Context indicators.** Indicators that are necessary to understand the context of the intervention and its relevance to the problems at stake and the external influences. These indicators are generally based on statistical information or administrative data.

These indicators are used to answer the evaluation questions, compare results between them or with some benchmark, assess the evaluation criteria and deliver a judgment. For example,:

• an intervention is efficient if its unit cost is inferior to the cost of other interventions that produce similar results.

• an intervention is effective if it reaches some specific and measurable objectives.

• an intervention is relevant if it can influence the initial problem in a significant way.

Analysis of indicators also helps explain to the taxpayers how well their money has been spent. Not all data and indicators are needed in all evaluations and sometimes it is difficult to identify from the planning stage all the necessary data and indicators. This is because some of them may emerge in the design of the evaluation study or during its implementation. However, in the plan, it is necessary to identify the basic data needed for each evaluation as well as understand the possible limitations in its availability.

Then, each of the planned evaluations need to matched with a set of data. At this stage, data sources and their availability must be checked. The single variables or indicators do not necessarily be identified at this stage as they will be developed by the evaluator during the design stage.

In some cases, when the evaluation is relatively simple and the evaluation questions are straightforward (e.g., efficiency of the implementation, or investigation on the preferences of the beneficiaries), the data needed is intuitive and does not require complex thinking. By contrast, a more careful appraisal of the needed data is required when the evaluation includes several different questions and evaluation criteria and as well a complex estimation of the impact. In this case, two main activities allow the matching between evaluation and data:

1. The association of the evaluation questions with the related indicators and, consequently, with the data necessary to calculate the indicators.

2. The association of the adopted methodology to data necessary to its implementation.

An example of the first method is presented in Table 8. The example relates to the design of an impact evaluation of an intervention aiming to promote start-ups. To each evaluation question, one or more indicators are associated; indicators require specific data, and this data is available in different specific sources. In the evaluation plan, the check may be less detailed and single indicators are not specified at this stage (they will be specified only in the following designing stage); however, the logic of this exercise and a first generic identification of the necessary indicators can be very useful in identifying needed data for each evaluation.

TABLE 8 **Relationships between the evaluation questions and necessary data: example from an impact evaluation on an intervention supporting the start–ups**

Questions	Possible Indicators	Necessary data	Source	Information Unit
Has the intervention produced additional effects on the creation of start-ups?	- Growth rate of start-ups - Birth rate of the start-ups before and after interventions	- Number of start-ups - Information on the characteristics of the start-ups	- Administrative: Register of enterprises	Territory covered by the intervention
Has the intervention produced positive effects on the growth and the development of the start-ups?	- Rate of survival after several years - Rate of growth of the investments of the start-ups - Financial and economic performance of the start-ups	- Total investments of the start-ups - Panel data on the survival of the start-ups - Turn-over and EBITDA of the start-ups	- Primary: Survey on start-ups, - Administrative: Register of enterprises - Administrative or private database with balance-sheets	Enterprise
Has the intervention generated positive effects on the human capital?	- Newly created employment per level of education	- Employment (total and new in start-up) by age, education, etc.	- Primary: Survey on start-ups - Administrative database with labor contracts	Enterprise

Questions	Possible Indicators	Necessary data	Source	Information Unit
Has the intervention increased the innovation capacity of the start-ups?	- Patents of the start-ups on the turn-over by sectors - Quota of product and process innovations on total innovations by sector - Investments in R&D of the start-ups	- Number of patents produced by the start-ups - Number and type of innovations introduced in the last years - R&D expenditure and employees - Presence of R&D equipment in the start-ups	- Primary: Survey on start-ups, - Administrative: register of patents	Enterprise
Has the intervention favored access to credit the start-ups?	- Average rate of interest paid by the start-ups - Rate of debit with financial institutions on the total debit - Presence of venture capital	- Interest paid by the start-ups - Short-term and long-term debit with financial institutions - Weight of private and public venture capital	- Primary: Survey on start-ups - Administrative: monitoring data on the conditions of the start-ups (applications)	Enterprise
Has the intervention produced multiplication and spin-off effects on the territory?	- Variation of the employment in the territory of the intervention - Variation in the number of companies - Specialization indicators in the sectors affected by the start-ups	- Number of employed people in total and in the sectors affected by the start-ups - Number of companies - Production by sector in the territory	- Official statistics and House of commerce data	Territory covered by the intervention

Source: Developed by Authors

The example in Table 8 is mainly oriented to produce a counterfactual and quantitative evaluation. Nevertheless, qualitative analyses follow the same logic from the evaluation questions to the needed data. In this case, it is important to pay attention to different data providers; for example, in a case study or in a survey, different subjects may be interviewed (beneficiaries, implementers, potential final beneficiaries, involved administrations or local organizations, etc.). The triangulation of sources is often crucial in these analyses, and this may affect time and costs of the data collection.

The possible mix of sources should be considered, too. For instance, administrative data may provide the universe for defining a sample for a survey; or monitoring data may be compared with administrative and official statistical data. These combinations often raise problems of comparison between sources, but they are very frequent and need to be foreseen in the plan.

The second technique (the association of the methodological approach with data) derives from the type of information that a certain methodological approach generally requires. Clearly, data needs may vary for each single evaluation, but some combinations are generally valid and may help in identifying the data to use. The most frequent methodological approaches and data sources are summarized in the Table 9 below.

TABLE 9 **The relationship between methodological approach and necessary data**

Methodological approaches	Frequent sources of data
Procedural analyses (procedural journey tracing, Input-Process-Output (IPO), data flow diagrams)	- Administrative data on the functioning of procedures (time, effort, subjects involved, etc.). This data is often available in the management IT system or must be collected by interviews - Projects and administrative documentation
Market and socio-economic analyses of beneficiaries and final beneficiaries	- Surveys on beneficiaries and users - Official statistical data to analyze context and to compare to programme outputs
Theory-based (Theory of change, process tracing, contribution analysis, impact pathways, Realist evaluation)	Often a mix of other sources is used: - Literature review - Projects documentation - Surveys and interviews to analyze behaviors of the actors - Administrative or official statistical data to measure effects of the intervention
Counterfactual – experimental (Random Control Trials (RCTs), Quasi experiments, Natural experiments)z	In general, there are two alternative approaches as regards the data needed: 1. Administrative data (less expensive, based on common administrative definitions, limited in the available variables) 2. Survey (more expensive, less precise in definitions, more extensive and flexible in variables on interviewees) 3. Official statistics, mainly when available at micro-level (individuals or territories);
Cost-Benefit Analysis	- Project data (technical parameters and estimated effects) - Official statistical data to have a set of context indicators
Case studies (Ethnography, QCA, within-case analysis, network analysis)	- Direct in-depth interviews - Surveys - Projects documentation
Statistical (Statistical modelling, Longitudinal studies, econometrics)	- Literature review - Official statistical data and administrative data
Synthesis studies (Meta-analysis, narrative synthesis, realist-based synthesis)	- Literature review - Other evaluations

In many cases and according to your experience, only one method is sufficient to identify the main data source to use; however, the two methods can be used together to reinforce data identification.

After having identified the main data to use and its sources, it is necessary to verify its availability for the evaluations. As already mentioned, at this stage it is not requested to enter in the detail of any single variable or data definition, but it is useful to know when and how data can be used. For this purpose, some simple criteria may be applied:

- *Accessibility:* Is the data public? Or does it have to be purchased? Or does it have to be formally requested from a specific owner? In the former case, does the MA make the request or can the evaluator do it?

- *Coverage:* Does the data cover the necessary lag time of the analysis? Does the data cover the territorial unit (e.g., province, region, country) required by the analysis?

- *Aggregation:* Is the data available at the necessary level of aggregation? (e.g.„ individual or aggregated data)

- *Protection:* is there a need for safety measures due to data protection rules? (e.g., monitoring data when both collected and used should be protected, as well as many administrative data)

Answering the above questions allows understanding of the general availability of the necessary data and the overall feasibility of the evaluation. Later, in the definition of the terms of references or in the evaluation design by the evaluator, a more detailed screening will result in detailed information on the data to collect and the limits of the analysis depending on the data characteristics.

4.4.2. Possible actions to improve data accessibility

As requested by the CPR 1060/2021, the MA should play an active role to ensure the data necessary for the evaluations.

First, playing an active role means making monitoring data easily accessible to the evaluator and researchers. The EC and some member states (e.g., Italy) make monitoring data available on their websites (https://cohesiondata.ec.europa.eu/ and https://opencoesione.gov.it/it/). In general, this data is stored in the IT system of the MA and can be easily retrieved.

Secondly, a possible improvement in data availability can derive from the National Institute of Statistics (INS) which could produce specific indicators at disaggregated territorial level for the cohesion policy analysis. In many cases, data for these indicators is available, but is not calculated in the traditional publications or sample of relevant surveys are not sufficiently numerous to cover regions or provinces; therefore, an agreement with the Institute and specific funding may reinforce the usable statistics.

Moreover, the MAs can contribute by providing key data for evaluations in the following ways:

•   Ensure a correct adoption of data protection rules

•   Stipulate agreements with administrative data owners. This simplifies and quickens the access to data of the evaluators and generates unique standard procedures for accessing to administrative data.

4.4.2.1. Facilitating agreements with administrative data owners15

Annex C - Promoting a clear and easy way to comply with data protection rules lists the main EU rules on data protection. The GDPR rules do not exclude the implementation of an evaluation; as a result, some basic procedural and operational steps must be respected in order to make the evaluation secure. For example, in a counterfactual evaluation, monitoring data (treated people) is combined with data from a public register (control group) and an external evaluator carries out the Impact Evaluation. In this case, here are the most important steps:

1. The MA comes to an agreement with administrations responsible for data (e.g., unemployment register, tax register, etc.) necessary to identify the control group and analyze treatment and control groups. The MA verifies that the use of this data complies with art. 6(1) e), or art. 6(4) and 5(1)b) of GDPR, if expressed consent was not collected.

2. The MA makes agreements with the other entities (data owners and the evaluator) to regulate the flow of information and the mutual responsibilities in accordance with GDPR rules. The service contract between the evaluator and the MA must include a specific clause on data protection; in the case of other administrations, a memorandum of understanding or specific national procedures can regulate data protection when implementing CIEs.

3. From the above agreement, the MA receives data in pseudo-anonymized form from the data-owner and transfers data for processing treated and control groups to the evaluator in compliance with the GDPR.

4. The storage of personal data of treatment and control groups only complies with the rules of data storage for the duration and purposes of the research and in accordance with basic security rules, as stated in the GDPR. These rules are observed by all the entities involved in the CIE.

National practices are consistent with GDPR, but sometimes there are serious obstacles to the use of micro-data. This may stem from the time required to adapt national rules and norms to the most recent GDPR, or it may be caused by different interpretations of GDPR in different national administrations, gold-plating, stricter GDPR rules in some cases, or misinterpretations. The MAs or the national authorities will need to take initiative to overcome these obstacles in the spirit of the GDPR and CPR Regulation.

---

15 This section is an adaptation of a chapter of: EC - DG Employment, 2021, Design and commissioning of counterfactual impact evaluations: a practical guide for ESF Managing Authorities, https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8426&furtherPubs=yes

4.5. DURATION AND BUDGET OF AN EVALUATION

The final characteristics of the planned evaluations concern their expected duration and the estimated cost. These elements are necessary to correctly schedule the evaluations across the period and allocate the necessary resources to the budget for evaluations. This section proposes some methods to address these issues.

4.5.1. Gathering information on market conditions

Having a dialogue with the market is important. Knowing and discussing with experts and private companies who will carry out the evaluations facilitates the understanding of the market conditions, the existing skills and experience, the possible options in the implementation of the evaluations.

In addition, a good dissemination of information can improve the quality of the tenders and open the participation to the calls for tender. It is important that these relationships are transparent and do not instigate conflicts of interest. To this aim, the meetings must be open, minutes of the meetings must be public, and discussions must focus on the contents of the evaluations and the operational solutions to improve evaluation activities.

This dialogue can help the administration to understand the practical issues restricting evaluation, the organization of evaluators, as to support a better definition of the timing and costs of the evaluation. The estimate of these elements of the evaluation, however, also depends on how experienced are the public officials involved in the definition of the evaluation plan. If these officials don't have any experience in the field of the social research, it would be appropriate for them to be briefed on these topics with colleagues or external specialists that have such experience.

4.5.2. Estimating duration and budget of evaluations

It is noteworthy that the lack of coherence between the evaluation questions and the available time and resources generally leads to poor evaluation results or a waste of resources.

Below are some suggested methods to estimate time, skills, and financial resources to implement an evaluation. These methods cannot be used in a mechanical way because each evaluation has its own complexity, but they provide a systematic approach to this issue.

4.5.2.1. Assessing the efforts needed and duration

The working effort and the duration of an evaluation are important elements to consider.

The effort incurred by an evaluation can be roughly divided into five principal activities[16] which are necessary in almost all the evaluations:

- **Evaluation design** involves defining the detailed methodology of the evaluation. This includes describing evaluation questions, how to collect, organize and analyze the needed information including the design of the instruments for data collection and analysis, as well as how to report the results.

- **Coordination** involves coordinating the evaluation team, attending the meetings with the evaluation commissioner, presenting results to the commissioner, the MC and other actors, as well as ensuring the quality of the evaluation. The administrative part of the evaluation is also covered here, including management of experts and payments.

- **Data collection** involves all the actions needed for collecting data from the sources identified in the methodology (documentation, surveys, interviews, and other types of field work, etc.) and storing data to be easily analyzed.

- **Data analysis** involves the implementation of the analyses, as to produce clear evidence and findings necessary to address the evaluation questions.

- **Reporting** involves the activity necessary to communicate results and includes drafting the reports, providing visualization and summaries of the results.

In case of complex or mix-methods evaluations, data collection and data analysis may involve many different activities; in this case, estimation of the evaluation duration and efforts should consider the combination of the different methods and the possibility of implementing them in parallel or not.

---

16 In some complex or participative evaluations also "evaluation design" can become an important activity and must be included in the budget estimation.

Table 10 proposes some methods to estimate the duration of the evaluation (from the signing of the contract to the approval of the final report and the effort (measured in working days) of some of the basic evaluation activities.

TABLE 10 Methods to estimate duration of an evaluation

	Duration	Effort
Evaluation Design	- It is completed in the inception phase, and includes the methodological and operational planning of the evaluation (methodological approach, analytical techniques, necessary data and the instruments to collect and store data quality assurance methods, the effective team and its organization) - It is presented in the Inception report. - Normally this activity lasts from six to eight weeks.	- Design generally needs not less than 15-20 working days, with preparation varying in complexity depending on the quality of the initial offer, and the scale and the difficulty of the evaluation, and the availability of various stakeholders involved.
Coordination	- Covers all the activities of the study and lasts for the entire duration of the study.	- It is generally not less than 10-15 working days (w.d) and around 10-15% of the overall w.d. or personnel cost.
Data collection	- **Surveys like CAT and CAWI (Computer-assisted telephone interviewing and Computer Assisted Web Interviewing, respectively)** can be requested to specialized companies. In general, they take from one to three months including the preparation of the addresses, recalls, etc.	- The effort can be estimated according to the number and length of interviews - Market costs can be easily found by specialized companies
	- **Direct interviews** for in-depth analyses or case studies take around one week from the request to the implementation	- The effort can be estimated in relation to the number of interviews and assuming a workload between 0.5-1 w.d. each (including preparation, arranging the appointment, carrying out and resuming the interview)
	- **Statistical data** from official sources is generally easy to collect and the creation of a database of reduced complexity can require two to three weeks, if data is available (extracting, storing, adjusting for the study, etc.), or even more than this, if data has to be requested to the responsible authorities.	- Due to the relatively easy accessibility of this data and the absence of significant idle time, in general, the effort is quite similar to the duration.
	- **Administrative data** often requires a significant amount of time to be released and made available; hence, duration must be estimated considering the time when the data is requested. Data preparation may also require a certain time because this data is not organized for statistical aims.	- The effort for requesting and preparing this data may be significant. Request can take time and need several interactions with the data owner if this is not used to delivering data for research. Preparation is often complex because this data must be reorganized and interpreted for the aims of the study. For instance, a database for a counterfactual impact evaluation can take an effort equivalent to 30-40 w.d. or more.

	Duration	Effort
Data analysis	- This activity can be classified with some difficulty in few specific categories and significantly differ according to the methodology, the type of data, the number of evaluation questions, etc. This means that an estimation of the duration must be contextualized to the evaluation. - The duration can be based in part on the data collection activities, because each collected data may require a specific analysis. - In general, evaluations with multi-source data and multi-method analysis require at least two months for the results.	- The same precautions expressed for the duration are valid for the effort. - In general, to have an approximate measure, the effort for data analysis can require a number of working days similar to that specified in the duration (60 w.d. as per the example in the left cell).
Reporting	- In general, this activity includes at least the inception, the interim and the final reports. Hence, the durations of these reports must be added up. - The duration of the reporting can vary between two weeks to two months depending on the type and the complexity of the report, the team available and the requested accuracy (interim, final, etc.). - The duration of the reporting, the time needed to review and finalize the individual reports must be estimated. In general, the commissioner should be able to send back its comment in two weeks.	- The effort for reporting should not significantly exceed its duration (see left cell). Clearly, the time needed to provide comments to the evaluators should not be included in the effort. - It must be also considered that, in some cases, the final report might benefit from the interim report, and it will focus only on final results.

Source: Developed by Authors

4.5.2.2. Identifying the requested skills and expertise

To prepare terms of reference and select the evaluator (see also chapter 6), it is important to know the type of team needed for carrying out the evaluation. Also, in this case, it is impossible to have a unique formula to identify the best team of all the evaluations, but some general guidelines can be provided:

- Evaluation requires to mix expertise and experts.

- When the complexity of the evaluation is high and the use of different techniques is demanded, additional expertise in data elaboration and statistics techniques may be necessary (econometrics, modelling, data mining, etc.).

- Similarly, when the field work is particularly intense, expertise in participative methods, case studies may be useful.

Evaluations need a mix of the levels of experience and different roles in the team. Consequently, evaluation teams generally include a senior coordinator, senior experts, and junior experts.

4.5.2.3. Estimating the possible cost

After the preceding clarifications are made, the next phase is the estimation of the evaluation budget. The evaluation budget usually includes the following costs needed to carry out the above activities:

- **Personnel** (e.g., evaluator(s), research assistant, support staff, etc.) per day or lump sum.

- **Travel** (transportation, per diem, travel mobilization expenses, consider class of travel).

- **Supplies**, equipment, and communication costs.

- **Workshops or other events** (design, findings verification, utilization, etc.)

In the absence of other references, the cost of the staff can be calculated in relation to the ESF+ fees adopted in the country for senior and junior trainers or experts. These fees are generally comparable to the costs of a junior and senior researcher and are a useful parameter to take into consideration. A simple market analysis and some interviews with a few researchers may help to specify the expected cost for personnel and other voices.

TABLE 11 Structure of the main costs of an evaluation

Main Activities	Evaluation team			Other Costs	Notes
	Program Manager	Senior experts	Junior experts	(materials, supplies, travel, etc.)	
Planning and coordination					The planning process includes the feasibility of the evaluation, its organization and the finalization of the methodology in agreement with the MA. In this phase enough time needs to be dedicated to the analysis of potential data gaps. The coordination covers the entire duration of the evaluation and includes the organization of the work and the interactions with the MA and other relevant stakeholders to finalize the evaluation questions.
Literature review					The literature review supports the outline of the theory of change, the identification of the variables to be used and the understanding of the context. It also allows refining the evaluation questions and exploiting past results and formulating hypotheses.
Data collection and preparation					The effort required for this activity varies widely depending on the data collection method(s). Original data collection via surveys can be time-consuming and expensive. Other methods may require an investment in technology (software or hardware) or finding an agreement with the data owner. Costs can be reduced by involving data owners in data preparation.
Data Analysis					Data analysis requires advanced skills and experience, and its duration is influenced by the method used, the analyses required and data quality.
Report(s) Preparation					The effort for this activity varies based on the number and type of reports and other required communication tools. Costs can include printing and graphic design in addition to preparation time.
Follow-Up Meetings					Follow-up meetings and other similar activities are an important step to disseminate findings. Miscellaneous costs could include space rental and food.
Total					

Source: EC - DG Employment, 2021, Design and commissioning of counterfactual impact evaluations: a practical guide for ESF MAs, https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8426&furtherPubs=yes

A distinction needs to be made between the evaluation of routine interventions, where expenditure is generally lower and innovative or pilot actions, for which the collection of a relatively high amount of data, the use of new data sources and the involvement of new stakeholders may justify higher expenditure.

However, this is not a rule and must be seen in the context of the evaluation questions, whose complexity and number may require lesser or greater effort.

A guidance document issued by the Commission[17] provides an indication of the amount needed for the evaluation of a program. This states that large-scale and

---

17   European Commission (2013) EVALSED: The resource for the evaluation of socio-economic development-

routine programs should dedicate no more than 1% of their budget to the evaluation; on the other hand, innovative or pilot initiatives may commit resources up to 10% of their budget. This high amount must be calculated only on the most significant, innovative interventions and interventions that have not been evaluated before (e.g., measures to prevent climate change or childhood support in this programming period) and not on the entire OP. Consequently, it is possible to presume that

resources for the evaluation could range from 1% to 2% of the total resources of the OP. However, in budgeting the evaluation, it is good practice to start from the information needs, and not from a theoretical budget to spend. It is expected that the budget will vary widely, subject to the abovementioned elements, but also considering the type of the evaluations to be carried out, in terms of scope and complexity, as well as resources involved.

## 4.6. CHECKING THE EVALUATION FEASIBILITY

Table 12 proposes a simple checklist to verify whether all the main aspects of a single evaluation (that need to be included in the plan) have been considered and, in general, if the evaluation is feasible. Questions are formulated in a way that positive answers ('Yes') mean

that the evaluation is feasible, while negative answers ('No') indicates that something in the planning has to be revised or that the evaluation could run some risks in terms of quality and feasibility.

TABLE 12 **Checklist of the evaluation feasibility**

Issue	Check	Yes /No
Evaluation questions	- Are the questions real evaluation questions, or anyhow useful to address evaluative issues?	
	- Is the number of key questions adequate and not too many (around 5-7 key questions)?	
	- Are the questions understandable and answerable?	
Existing knowledge	- Are responses to the questions not available in reliable past evaluations or relevant studies?	
Unit of analysis	- Is the unit of analysis clearly identified and defined? (What part of the OP, how many OP resources involved, type of intervention, etc.)	
Timing	- Is the timing of the evaluation coherent with the evaluation questions?	
	- Is the timing of the evaluation coherent with the expected progress of the interventions?	
Stakeholders	- If stakeholders are influential on the definition and use of the evaluation, have they contributed to the evaluation questions?	
Methodology	- Is the methodological approach well identified and clear?	
	- Is the methodological approach coherent with and cover all the evaluation questions (e.g., what? why? how?)	
	- Is the methodological approach feasible with the existing skills in the market?	
	- Is the methodological approach feasible with the available resources and time? (In relation to data collection, field work, etc.)	
Data availability	- Have the necessary data for the evaluation been identified?	
	- Are the necessary data for the evaluation available and accessible?	
	- Are data protection rules and other CPR requirements respected?	
	- If some data must be collected by the evaluator, is direct collection possible?	
	- Has the MA activated agreements or interactions with other administrations or data owners to make data available?	
Resources	- Are the costs of the evaluation adequately estimated and do they consent to adequately carry out the study?	
	- Are the costs of the evaluation reasonable and affordable in comparison to the entire budget for evaluation?	
Internal capacity	- Are the needed expertise and capacity to supervise the quality of the evaluation available in the MAs?	
	- If additional expertise is needed to support the MA in the implementation of the evaluation, can it be acquired?	

Source: Developed by Authors

EVALUATION FRAMEWORK

5
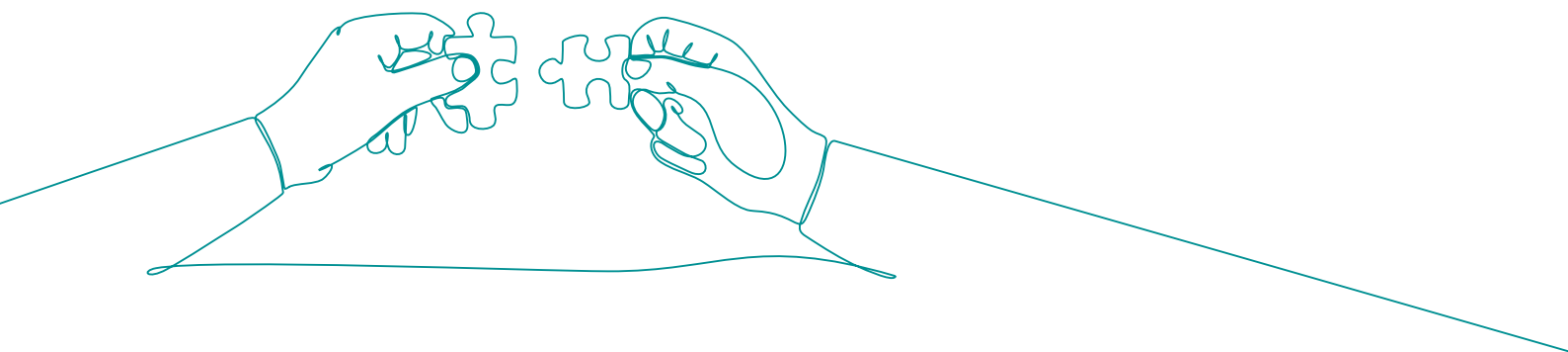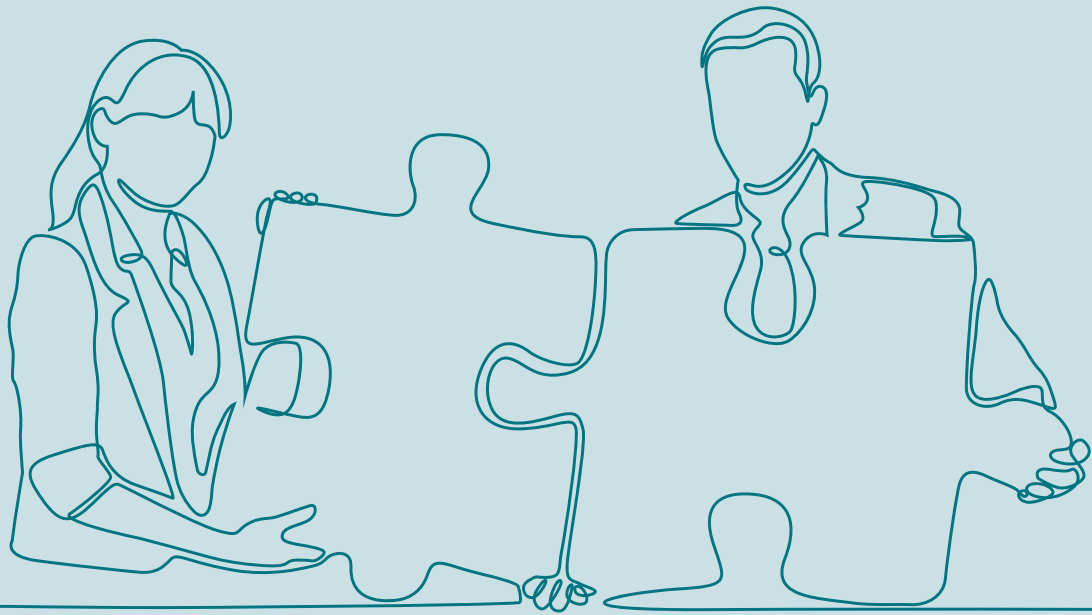
This chapter aims to define the arrangements needed for an effective management of evaluation processes, to produce high quality evaluation results and promote their use. As a result, it defines the evaluation governance (or "evaluation framework", as called in the working document of DG Regio) for the preparation and implementation of 2021-2027 evaluation plans in Romania, by both individual MAs and the Evaluation Central Unit.

The evaluation plans, as requested by the EC working document, should include a presentation of the basic setup of the evaluation system and the instruments needed to shape and reinforce evaluation quality.

In line with the Commission's recommendations and the template of the evaluation plan presented in Annex A – Template of the evaluation plan, the evaluation framework covers the following aspects:

- **National governance of the evaluation** (MA institutional arrangements, coordination inside and outside the MA's administration, working groups, etc.)

- **Coordination and organization of the evaluation at the OP level** (This section introduces the organization of the evaluation at the OP level: the role and responsibility of the different actors)

- **Involvement of the stakeholders in the evaluation plan** (who and how has been involved during the preparation of the plan, instruments planned to involve stakeholders during the plan's implementation)

- **Strategic approach to the evaluation** (how the planned evaluations have been identified and what are the most important and strategic elements of the plan)

- **Evaluation implementation and quality** (launching calls for tender to select evaluators, issues related to planning evaluations, preparing ToRs and defining procurement arrangements; quality of the evaluation process, main requirements to ensure quality, possible risks)

- **Dissemination and use of the findings** (multi-target communication and involvement of the stakeholders, basic tools and venues to disseminate findings, use of the policy recommendations and their implementation).

- **Evaluation capacity building** (inside and outside the administration, training, relationships with the market and the academy, involvement of the stakeholders)

- **Available resources and timing of the evaluation** (timetable of the evaluations and their estimated costs)

Some of these aspects were already examined in the previous sections and more precisely: "Strategic approach to the evaluation" is explained along all the previous chapter 4 and "Available resources and timing of the evaluation" is specifically discussed in section 4.5. Consequently, the following sections will examine the remaining aspects of the evaluation framework.

5.1. GOVERNANCE OF THE EVALUATION SYSTEM OF ESIF

The governance of the evaluation system consists of institutional arrangements at both central level and individual OP level. The central level carries out national evaluations, such as evaluations of the PA and horizontal or cross-programme evaluations, coordinates the evaluations of the OPs and promotes a general and homogeneous reinforcement of the evaluation capacity. The OP level evaluates the single OP, developing sectoral and territorial approaches, and promotes the dissemination and use of findings among the actors of the OP.

The governance of the evaluation system must guarantee some basic functions:

- Defining the evaluation plans of the OPs

- Carrying out evaluations respecting timetable and main contents defined in the evaluation plan

- Ensuring the quality of the evaluations

- Promoting dissemination and use of the findings of the evaluations

- Reinforcing evaluation capacity within the administrations and among the stakeholders

- Coordinating the evaluations of the different OPs with national/horizontal evaluations and PA's evaluation.

The institutional set-up involves creation of individual evaluation units in each MA, giving the role to coordinate and guide the units to the MEIP ECU. The roles of the different entities complement each other, aiming towards a more hands-on evaluation process. Also the intention is to be closer to decision-makers to better respond to the information needs of the different actors involved (MAs, IBs, policy makers, implementers, etc.).

The evaluation governance includes the following main actors:

- The **Monitoring Committee** at OP level and the Coordination Committee for the Management of the Partnership Agreement (CCMAP) at PA level, as requested by the EC Regulations, must approve the Evaluation Plan and examine *"the progress made in carrying out evaluations, syntheses of evaluations and any follow-up given to findings"* art.40 (e) 1060/2021.

- **Evaluation Central Unit (ECU)** is responsible for the evaluation of the Partnership Agreement (PA) and cross-thematic or territorially focused evaluations. ECU also plays a pivotal role in ensuring coordination across the system, both for the evaluation process and capacity building.

- The **evaluation units at OP level**, support the MA in its responsibilities for OP evaluations. They are composed by experts in evaluation and research associated to the staff of the MA with the specific and unique assignment of supporting evaluation and monitoring processes.

- The **Evaluation Steering Committee (ESC) at PA and, respectively, at OP level** plays both a methodological function and a key role in ensuring that key actors are consulted and represented in the planning and implementation of evaluation and are aware of evaluation results. To this end, it will generally include representatives of other MA evaluation units and ECU, key stakeholders with a role in policy-making for the areas covered by the PA/OP (e.g., line ministries), data suppliers (e.g., the National Institute for Statistics), representatives of the academia or social partners. The ESC should function in a flexible manner as to accommodate the arrangements made for each OP or group of OPs (e.g., in case of ETC programmes) but it will meet at key moments of the evaluation, such as the completion of the terms of reference, the discussion of the evaluation methodology, the partial and complete results and the use of the conclusions and recommendations of the evaluation.

- The **Evaluation Scientific Committee** (SC) can be set-up at evaluation level, to support ECU and the MA and the evaluation unit during the implementation of the evaluations. This should be composed by two to three senior experts an will help to ensure that evaluations meet the quality standards in terms of methodological choices, reliability of findings and strength of recommendations. The experts may be appointed for each evaluation and contracted by the evaluators (at the recommendation and/or with the approval of ECU/ the MA). Alternatively, a mechanism can be set up at PA/OP level, to finance the costs associated with contracting these experts when needed. The SC can be used only for complex evaluations, where the evaluation unit does not have enough capacity to ensure a quality evaluation process.

- The set-up and institutionalization of the national **Evaluation Working Group** including the ECU and the evaluation units; this group plays an active role in making sure there is a coordinated approach to evaluation and promotes general initiatives for increasing evaluation capacity. ECU is the leader of this group, by preparing the agenda of the meetings, providing the technical assistance, and facilitating capacity building and exchange of experiences.

- While it is an informal mechanism, the **Evaluation Network**, set up during the 2014-2020 programming period, will continue to function in the current period and will act as a forum for discussions and exchange of good practices and lessons learned. This includes both members of the contracting authorities, members of the academia and the evaluators.

Table 13 below reflects the roles of different actors in relation to the main activities to plan and implement evaluations.

TABLE 13 Main roles in the governance of the evaluation of ESIF

Action/ Process	National level	OP level
Approving evaluation plans and examining progress in evaluations	**CCMAP** Approves the evaluation plan and examines progress in carrying out evaluations	**Monitoring Committee** Approves the evaluation plan and examines *"the progress made in carrying out evaluations, syntheses of evaluations and any follow-up given to findings"* art.40 (e) 1060/2021
Implementing evaluations	**ECU in MEIP -** Implements PA evaluation and other horizontal evaluations	**MA Evaluation Units -** Implements OP level evaluations as to provide technical support to the MA on monitoring and evaluation
Methodological support and ensuring institutional representativeness	**Evaluation Steering Committee (at PA level) -** discusses the evaluation methodology and the partial/ final results of evaluations of the PA and how to improve the use of evaluation results (also selects the recommendations to be implemented)	**Evaluation Steering Committee (at OP level) -** discusses the evaluation methodology and the partial/ final results of evaluations and how to improve the use of evaluation results (also selects the recommendations to be implemented)

Action/ Process	National level	OP level
Providing technical support for evaluation	**Evaluation Scientific Committee** (per each evaluation)	
	It is composed by two to three senior experts (members of the academia or senior sector or evaluation experts), who support the MA, evaluation unit/ ECU and the evaluator during the implementation of a single evaluation.	
	The support is discretionary and can be used for large or complex evaluations, where the evaluation unit does not have enough capacity to ensure a quality evaluation process. The experts can be hired directly by the MA/ECU or through the evaluator, including this obligation in the ToR and selecting them with the agreement of the MA/ECU. The required effort should not exceed three to four days per expert in each evaluation	
Methodological coordination of evaluation activities at national and OP level	**ECU** to provide methodological coordination and support for the MAs evaluation units charged for the evaluation of ESIF, to support capacity building at the level of the Evaluation units and facilitate a coordinated approach to evaluation.	
	ECU will participate in relevant evaluation networks of the European Commission and relevant know-how will be passed on to the Evaluation Units, through the mechanisms defined, as per the planned evaluation governance.	
	Evaluation Working group (chaired by ECU)	
	This working group/network chaired by ECU and including all the Evaluation Units aims to ensure coordination and methodological support for the evaluation activities of the OPs and is organized and activated at the beginning of the programming period.	
	The institutionalization of the Evaluation Working Group (in the PA or in some national act) will reinforce the functioning and power of both the Group and ECU. Funding for the functioning of the Evaluation Group needs to be ensured.	

Source: Developed by Authors

Some aspects are crucial for the effective functioning of the evaluation system. These pertain to both the MAs and the ECU but are also related to the relations between stakeholders.

Table 14 below presents in more detail the roles of the MA Evaluation Unit and the Central Evaluation Unit in MEIP. As mentioned, close cooperation is ensured throughout the programming period for an effective evaluation process and use of the evaluation results.

TABLE 14 Main roles of the MA Evaluation Units and ECU

Field	MA Evaluation Unit	ECU18
Evaluation Plans	- Identifies and engages relevant stakeholders for the program evaluation - Identifies program evaluation questions - Prepares the evaluation plan of the OP in line with the guidance of ECU - Revises and updates the plans as needed	- Prepare guidance for elaborating the plans (Template for evaluation plans, checklists for quality verification, etc.) - Review of the evaluation plans and formulate recommendations for improvement
Evaluation procurement	- Prepares ToRs based on national guidance - Support the MA in the tender process to identify external evaluators	- Guidance for optimizing the procurement process (templates for ToRs, guidance on selection criteria, etc.)
Evaluation management	- Ensures the management of the OP evaluations and relevant studies	- Provides general how-to-do guidance for an effective management of evaluations (verification, schedule, etc.)

---

18 Fulfill the same role as an MA Evaluation Unit for the evaluation at PA level

Field	MA Evaluation Unit	ECU18
Evaluation Reporting	- Prepares/Verifies the program evaluation reports with the help of the checklists prepared by ECU - Presents evaluation findings to the MCs	- Provides guidance on how to prepare/assess Evaluation Reports - Publishes the evaluation reports on the evaluare-structurale.ro website
PA and horizontal or thematic evaluations, studies	- Provides necessary information and includes in the OP evaluations horizontal and comparable elements requested by ECU	- Carries out PA/macro level and thematic evaluations - Delivers studies and reports on topics and performance of the OPs
Synthesis reports	- Provides information on the evaluations carried out and the main findings/ recommendations (according to the guidance of ECU)	- Supports preparation of consolidated reports on fund implementation (by providing information on evaluations carried out under all programs)
Evaluation Working Group	- Attends Evaluation WG meetings	- Coordinates the activities of the National Evaluation WG, - Hosts the secretariat of the Evaluation WG - Organizes the meetings of the OP working group
Evaluation Network	- Voluntary participation of staff in the meetings organized in the framework of the network	- Chairs the network and organizes events - Publishes relevant information on the Network activity on the website
Skills development	- Participates in trainings organized by MA/ECU/EC or other entities	- Coordinates and manages the training process
Communication of evaluation results	- Communicates the results of the OP evaluations to the key stakeholders (including decision-makers) and the public	- Provides guidance on how to improve communication of results - Communicates results of PA/cross-OP/meta-evaluations/studies, etc. - Manages www.evaluare-sructurale.ro website
Connecting to European Practice	- Receive and use the information shared by the European Networks of the European Commission via Evaluation Central Unit	- Represents Romania in the Evaluation Networks of the European Commission and disseminates the information towards the MA evaluation units

Source: Developed by Authors

5.1.1. National governance of the evaluation

At the national level, the ECU plays a central role in the overall set-up and is in charge of both PA-level evaluation and ensuring the methodological coordination of the overall evaluation process and promoting capacity building at system level. To successfully carry out its evaluation tasks (related to the PA), ECU is supported by both a Steering Committee and a Scientific Committee, that offer technical support for all evaluations. At a higher level, the Coordination Committee for the Partnership Agreement approves plans for evaluations, while also supervising the use of evaluation results.

The National Evaluation Working Group also plays an active role in coordinating methodological efforts at national level. This is composed of representatives of all MAs evaluation units, as well as ECU, which plays a leading role. The coordination efforts are key in creating consistent practices across the board and in sharing good evaluation practices, as well as providing ways to give adequate guidance and support.

FIGURE 6 Set-up of the evaluation system at national level

Coordination Committee for the Management of the Partnership Agreement
Approves PA evaluation plan, examines evaluation results and is informed about the use of recommendations

Steering Committee (at PA level)
Discusses the evaluation methodology and results, and how to use evaluation results and recommendations

Evaluation Central Unit (ECU)
- Elaborates evaluation plan for the PA, contracts and implements evaluations
- Ensures methodological coordination of evaluation, provides guidance and facilitates capacity building

Evaluation of the PA and horizontal evaluations

National Evaluation Working Group
Includes members of all MAs and ECU. Supports coordination of evaluation activities for all OPs and promotes a unitary approach to evaluation, facilitates sharing of good practices

Scientific Committee
Technical support for each evaluation (if needed)

Evaluation Unit 1 — OP 1
Evaluation Unit 2 — OP 2
Evaluation Unit 3 — OP 3
Evaluation Unit 4 — OP 4

Source: Developed by Authors

The Evaluation Coordination Unit (ECU)

The ECU plays an important role in achieving successful functioning of the evaluation units, by ensuring coordination and providing guidance on overall processes. ECU will have a coordination, consultative, and guiding role, which involves focusing on harmonization, promoting training and overseeing the evaluation process at ESIF level.

ECU plays a key role in supporting OP evaluation in the following ways: coordination of the evaluation plans to avoid overlaps or misinterpretation of EU rules, guidance for carrying out the evaluation activities (e.g., preparation of evaluation plans, evaluation management, verification of the evaluation reports, etc.), provision of training, implementation of PA level evaluations and coordination of evaluations on horizontal themes. In particular, the following areas would benefit from this coordination:

- Training of the staff in the MA and in the Evaluation Units and ECU, in areas such as evaluation methods and techniques, methods for collecting evaluation questions, guidance for designing evaluation plans,

defining ToR and procurement aspects, identifying lessons learned, communication of evaluation results, use of evaluation findings, etc.

- Support for evaluation planning at OP level, through providing consistent and unitary guidance on evaluation plans, exchange of experiences and peer-reviews of the evaluation plans.

- Support for preparing the ToR of the calls for tender of the planned evaluations, guidance on selection criteria, etc.; support of the MA in the tender process to identify external evaluators.

- Support for improving communication of evaluation results, through guidance on disseminating results to different categories of audiences, as well as publishing evaluation results on www.evaluare-structurale.ro;

- Organization of events, such as meetings and working groups, to ensure a common approach to evaluation of Cohesion Policy Funds.

- Provision of support materials on evaluation, such as regulations, guidance documents, templates, etc.

The National Evaluation Working Group

The working group has a consultative role and aims at favoring a coordinated approach to evaluation at the national level. To this end, the Group is chaired by ECU and includes Evaluation Units of the OPs or representatives of the MAs.

The Working Group regularly meets monthly or every two months, according to a pre-defined calendar. Aspects to be discussed include: approach to planned evaluations, common themes or complementary interventions at the national and OP level, methods and tools for different types of evaluation, challenges and good practices, training needs or evaluation capacity initiatives for different actors.

Internal procedures should be defined for an adequate management of the working group. MAs nominate members of the Evaluation Unit to participate in the National Working Group and attend all meetings in the working group. Guidance documents issued further to the meetings must be made available to all participants.

A system to facilitate communication and availability of the documents prepared should also be designed. This will allow access to all guidance documents available and facilitate dissemination of relevant guidance and information throughout the system, also to the staff/stakeholders involved in/interested in evaluation, but who are not directly members of the Evaluation Working Group.
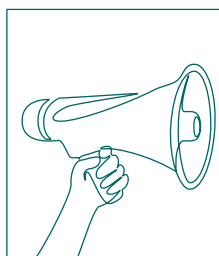
These main characteristics of the national governance of the ESIF evaluation system must be explained in the evaluation plan at the beginning of the section on the evaluation framework. It allows the readers to understand the context in which the evaluation plan is designed and how it is coordinated with the other plans. If agreed, all the evaluation plans may adopt the same presentation of the national governance.

5.1.2. Coordination and organization of the evaluation at the OP level

The institutional set-up at OP level (Figure 7) is similar, to some extent, to that defined at national level. The evaluation unit plays a central role in planning and implementing evaluations for the OP(s); to this end the Evaluation Unit will play a key role in providing the needed technical knowledge at the level of the MA, in relation to the other departments.

The MA activity in the field of evaluation is supported by the Evaluation Steering Committee, who can provide inputs to the preparation of the evaluations and discuss findings and facilitates use of evaluation results. The Evaluation Steering Committee appointed should actively involve stakeholders; in this framework, the MA together with members of the Monitoring Committee and other stakeholders, can discuss the objectives and the results of the evaluations in an in-depth and comprehensive way. The main decision forum is, however, represented by the Monitoring Committee, who approves the evaluation plans and discusses recommendations.

In its work, the MA may be supported by a Scientific Committee that may be summoned for all evaluations or only for more complex ones. Critical methodological support is provided by ECU, both directly (with the support of guidance and methodologies and by organizing activities aimed at capacity building) and in the framework of the National Evaluation Working Group.

The National Evaluation Working Group is composed of members of all MA evaluation units and ECU and and helps the MA units adopt a coordinated approach to evaluation, use methodologies and instruments, communicate results. It also helps ensure application of the guidance and support available and acts as an open forum for discussions and sharing experiences.

FIGURE 7 Set-up of the evaluation system at OP level
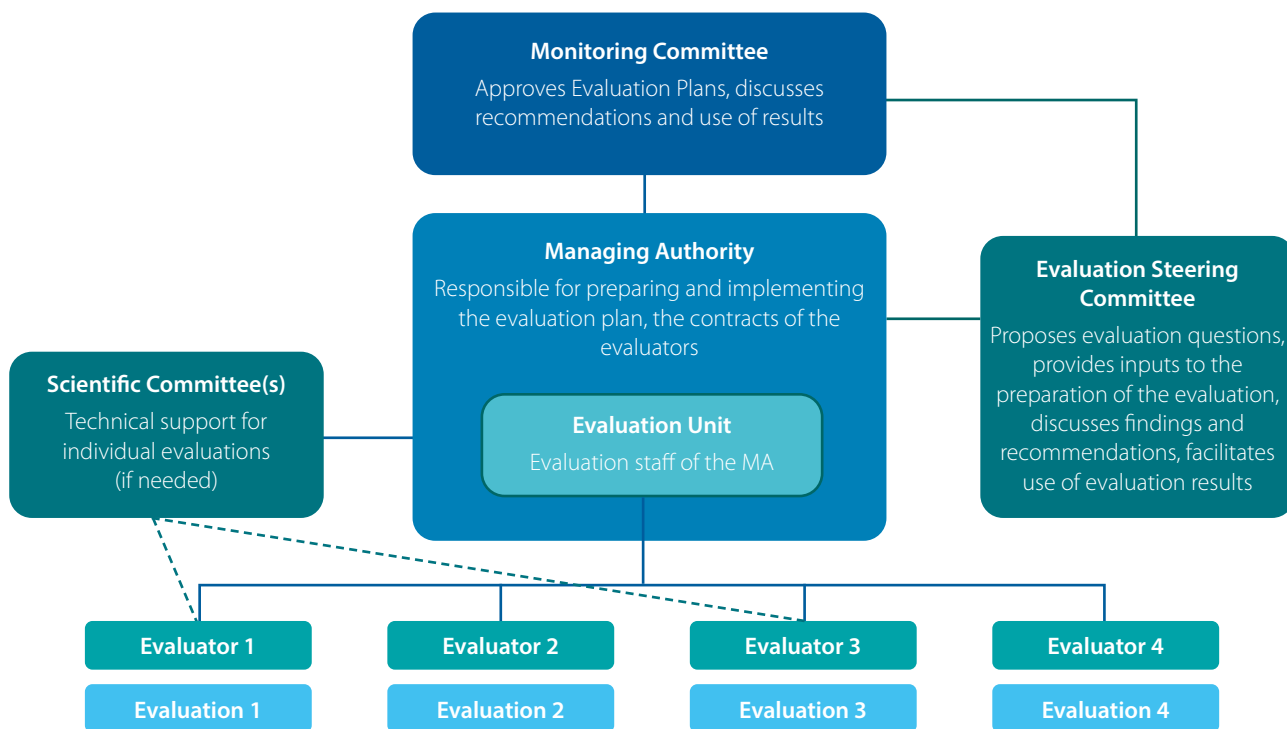
Source: Developed by Authors

The MA Evaluation Units

Adequate **staffing** of the Evaluation Units, in terms of number and skills, is also necessary. A minimum staff of two to three persons must be provided; the need for additional members may depend on the financial size and complexity of the OP concerned and of its evaluation plan. Evaluation officers need to have the knowledge on statistics, economics, sociology and research skills, experience in public administration, statistical analyses and participatory activities.

The **tasks** of the Evaluation Unit reflect the building blocks of the monitoring and evaluation system and are aligned with the overall mission of supporting the decision-making process. Tasks of the Evaluations Unit are defined per the Regulation of Organization and Functioning (ROF), setting up the tasks and responsibilities for all departments and directions in the institutions, as well as responsibilities of management staff. Internal procedures also need to be formulated to allow for a smooth implementation of the agreed tasks, as well as to facilitate sharing evaluation results and use of findings to support the decision-making process. Positioning the Evaluation Unit close to the MA general director and clearly defining its hierarchical relation with the other units (implementers, controllers, audit, monitoring, etc.) is also vital. Evaluation tasks may include:

- Designing, monitoring, and updating, when necessary, the multiannual Evaluation plan;
- Planning how evaluations must be carried out (procurement, resources, management arrangements, methods, and instruments, etc.).
- Delivering reviews of evaluations and meta-evaluations on themes of interest and in preparation of the evaluations.
- Drafting evaluation related documents (ToR, templates for Evaluation Reports, checklists, etc.).
- Defining criteria for quality assessment of the evaluation reports and processes.
- Setting up Evaluation Scientific Committees for each evaluation exercise (subject to complexity) and organizing and guiding their meetings.
- Supervising evaluations carried out by external contractors
- Carrying out in-house evaluations - if applicable.
- Identifying and engaging relevant stakeholders.
- Facilitating the meetings and functioning of the Evaluation Steering Committee for the OP, to assess

methodological aspects for the evaluations, analyze evaluation findings and recommendations and promote the use of evaluation results. Appropriate selection of the recommendations to be implemented should also be facilitated.

- Facilitating the participation of its staff in the meetings of the Evaluation Working Group, as an evaluation "forum" to discuss evaluation issues (evaluation questions, findings, etc.).

- Participating in Monitoring Committee meetings and promoting the use of the evaluation findings in the decision-making process.

- Ensuring communication of evaluation findings and recommendations to relevant stakeholders.

- Ensuring monitoring of the uptake of evaluation recommendations.

- Implementing or participating in activities to strengthen evaluation capacity (trainings, events, workshops, etc.).

> The governance and the organization of the evaluation at the OP level (or related to the OPs involved in the plan) must be described in the section 1.2 of the evaluation plan (according to the template in Annex A). The system should ensure the fundamental capacities mentioned above (involving stakeholders, implementing high-quality evaluations, using the results, etc.) and must clearly identify the responsibilities and the roles of the different actors and working groups.

5.2. INVOLVEMENT OF THE STAKEHOLDERS IN THE EVALUATION PLAN

Stakeholders will be involved in more stages of the preparation of the evaluation plan, right from the beginning of the process, when evaluation questions are formulated, up to the final stages when looking for ways to enhance the use of evaluation results. The Evaluation Working Group at PA level and the Steering Group at the OP level are important venues aiming to encourage involvement of the different groups of stakeholders throughout the entire evaluation process. These groups may be activated during the preparation of the plan to get a broad range of stakeholders involved; in addition, conducting specific meetings or surveys will bring about more targeted discussion with the usual group of stakeholders as well as find groups that will diversify the consultation pool.

As shown under section 4.1.1.4, involving the stakeholders in the design of evaluation questions, by encouraging them to propose their questions, is key to ensuring that the evaluation plan can respond to the needs of the different programme actors and for facilitating the future use of the evaluation results. Proper identification of the key stakeholders, together with definition and use of the right tools to interact with them are key to maximizing their inputs. Validation of the selected questions, after first prioritization and selection of questions by the MA based on the contributions received is also important before proceeding further with the evaluation plan.

The involvement of stakeholders is also essential for validating the methodological approach and ensuring data will be available throughout the evaluation process. Therefore, stakeholders should be consulted on the planned arrangements and informed on the expected data needs and timing for evaluations, as to facilitate their engagement throughout the entire process. While ensuring the consultation takes place and relevant inputs are collected and adequately reflected in the evaluation plan, the approach for involving stakeholders can be adapted to the needs and specificity of each programme, as to accommodate existing institutional and procedural arrangements.

The Evaluation Steering Committee at OP level – that includes members of all MAs evaluation units and the ECU, together with key actors involved in policy making, academia and data suppliers – is expected to play a major role in the process, both for ensuring that evaluations planned respond to the information needs and the methodologies foreseen are the most appropriate to the scope. Moreover, it provides continuity along the way, from the planning stage to the analysis of recommendations and further on, by facilitating the use of evaluation results, by supporting an adequate monitoring of the stage for implementing the recommendations.

Monitoring Committees and the CCMAP are additional forums where the stakeholders can analyze evaluation plans and give input to maximize the use of the results. Adequate preparation for the meetings – in terms of preparatory documents provided, adaptation of materials to fit the profile of the different audiences, initial guidance, timing of presentations – will help to capitalize on stakeholders' engagement and increase their interest and involvement in evaluation.

Stakeholders are also critical when trying to ensure timely availability of the necessary resources. Safeguarding the financial resources allocated for the evaluations is a key aspect that will be factored in the evaluation plan. Nonetheless, additional efforts will need to be made. The various stakeholders will contribute throughout the process with different resources that are not necessarily quantified in the initial planning, such as, for example, the human resources involved in data collection and reporting, the time allocated by such persons to participate to the interviews, focus-groups or surveys planned during the meeting etc.

In other words, stakeholders' involvement must be properly planned and provided for, in regards to each stage of the planning process. Such undertaking is not only needed to meet the EC's requirements on partnership, but to increase awareness of the overall evaluation process and optimize use of results.

The way this occurs during the preparation of the evaluation plan and how it will continue during the implementation of the evaluations must be described in the evaluation plan in section 1.3.

5.3. EVALUATION IMPLEMENTATION AND QUALITY

5.3.1. Planning and managing evaluations

Management of evaluations is a crucial area. Here are the aspects to be covered: staff allocation to different tasks, procedures established for launching evaluations, selection of the evaluators, verification of the evaluation reports, use of the evaluation results. All these activities must follow quality criteria and mitigation of risks to delivery high quality evaluations.

In this process, **respecting the timetable of the planned evaluations** is a complex task; it requires a precise forecast of the time it takes to select the evaluator, the right approximation of the duration of the evaluation (especially data collection and analysis) and estimation on the waiting time before outputs or outcomes will be available for analysis. In addition, the timetable must respect the EU regulatory provisions and other national obligations, if any.

A summary table (as proposed in the template of the plan in annex A) and GANTT (including time for public procurement and other preparatory activities) can be used for planning evaluations, considering their financial size, allotted time, and applicable tendering procedures. Sufficient time must be allowed for procurement, building on the lessons learned from the previous implementation period. Evaluations on complementary themes (at the level of the OP or across OPs), evaluations informing further studies/reports or larger scale evaluations must be carefully planned. The agreed upon calendar must be updated regularly, to reflect any changes of priorities or delays in implementation.

Evaluations must also be planned to allow **adequate management** by the MA Evaluation Units/ECU, avoiding overstretching of staff, lack of adequate skills or unsatisfactory involvement of stakeholders and users of the evaluation findings.

Ways to simplify procurement should be explored. The Guide proposes several suggestions for designing good technical specifications in the calls for tender for evaluators. Even so, it was difficult to define a single template of technical specification to use in many calls for tenders for evaluation; such model(s) or template(s) should be defined with legal support at the national or OP level. Consultation with the concerned authorities, such as National Agency for Public Procurement, is crucial to ensure agreement. Selection criteria should be also reconsidered, for increasing their consistency with EU evaluation principles and internal coherence (see also section 6).

5.3.2. Quality management strategy for evaluation

Several actions can be implemented to ensure a qualitative evaluation process. These relate to both institutional set-up for evaluation, instruments used, as well as other actions that can be taken to reinforce quality.

Missing arrangements for quality managements can incur a lot of risks. Among these, poor relevance of evaluations, significant delays in implementation of evaluations, poor management of evaluations, poor quality of evaluations reports, etc.

The Scientific Committee, as mentioned before, can be useful in quality assurance of evaluations when staff is still inexperienced or lacks the required skills for the evaluation. It can support the review process and the discussion of the deliverables, but it can also be involved throughout the process, to ensure the methodology and instruments selected are adequate for the theme/evaluation proposed. The SC should include two or three senior evaluation experts and/or members of the academia.

Standardized templates and instruments can also be used throughout the evaluation process to ensure quality. These can target both the planning and implementation of evaluations, but also the quality of the evaluation reports. Such templates include:

- A plan for managing evaluation (indicative steps & list for checking key milestones and assess potential risks)
- Template for evaluation report
- Checklist for quality evaluation process

These templates are included in the Annexes to the present Guide.

Other ways to reinforce quality of evaluation refer to the process itself: ensuring independence of the evaluators and ensuring the prerequisites for an optimal evaluation process (see above call for tenders – methodological aspects, selection criteria for the experts), ensuring adequate resources to the objectives, in-depth revision of the reports, involvement of the stakeholders in the preparation process and validation of the evaluations results. Table 15 below provides a 12-step summary of the most frequent risks encountered along the evaluation process and the possible mitigation measures to anticipate or counteract the risks.

TABLE 15 Risks and quality insurance in the evaluation process

		Most frequent Risks	Quality measures
1.	**Defining evaluation questions**	- Formal approaches - Only few persons making decisions on the questions - Too many or unclear questions	- Updating evaluation questions in the plan - Consulting stakeholders (e.g., policy makers and implementers etc.) - Identifying and selecting key questions
2.	**Who evaluates (internal or external evaluation)**	- Lack of independence - Lack of sufficient skills - Lack of organizational capacity in respect to time and resources available	- Using an external evaluator or introducing external experts in internal evaluations - Using a market study to verify available capacity and skills - Involving academics/experts in the evaluation team (with the help of the SC, ESC, as well as through dedicated events)
3.	**Writing terms of references**	- Unclear aims and use of the evaluation - Unclear or "not answerable" questions - Too many evaluation questions - Too detailed methodology and applicant has not freedom to propose - Not pertinent requested evaluation team - Resources and time not consistent with the request	- Using ToC to organize and prioritize evaluation questions - Focusing the evaluation on 5-7 key evaluation questions - Mixing evaluation and sectoral experts in the team - Having minimum experience on how to carry out an evaluation

		Most frequent Risks	Quality measures
4.	**Selecting the evaluators**	- Too many or too severe eligibility restrictions limit the market - Absence of quality criteria penalizes evaluation quality - Unqualified or little motivated evaluators - Lack of criteria or terms to verify quality and ability - Approaches that are too academic and thus incapable of addressing policy issues	- Ensuring open access, but with experience in evaluation and the examined sector - Lower weight of the price criterion in the total score - Using quality criteria - Involving experts in the selection committee, if internal skills are insufficient
5.	**Ensuring the evaluability**	- Lack of data, wrong questions or delays in the intervention may produce an incomplete and unserviceable evaluation	- Asking an evaluability assessment in the inception report and, if necessary, stopping the evaluation
6.	**Refining the methodology**	- Weak methodology or incapability of understanding evaluation questions often produce unreliable and biased evaluations	- Using the kick-off meeting and the inception report (after no more than one month) to clarify methodology issues - Not approving the inception report until it addresses the issues
7.	**Collecting information**	- Inadequate or insufficient information produces unreliable or weak general results - Too much information collected in comparison to the questions and the planned methodology produces confusion and ill-defined evaluations	Mitigations can depend on the selected methodology: - Counterfactual evaluations require enough cases and a well-identified control group - Theory-based approaches need data specifically related to the method and triangulation of information sources to avoid biased results - Surveys require well-structured questionnaires and efficient investigation (cases/costs) - Case studies require a selection of cases consistent with the aims (explorative, comparative, representative, etc.) - Mixed methods approaches need balanced integration and avoiding duplications or inconsistent linkages between different data
8.	**Analyzing information**	- Wrong or inaccurate use of the methodology or analysis techniques leads to inaccurate or unreliable findings - Lack of authority of and confidence in evaluation	- Verifying the implementation of the methodology before the end of the evaluation in an interim report - Requiring robustness or sensitivity analyses, where appropriate - Not accepting the first explanation, but requiring checking also alternative interpretations of results - Inquiring unexpected and indirect effects

		Most frequent Risks	Quality measures
9.	**Evaluating and answering the questions**	- The study is descriptive and is not an evaluation - The study does not provide judgements and related evidence, and does not answer the evaluation questions - Generalization of findings is unclear and not discussed - The findings are not translated in policy implications	- Defining evaluation criteria in a clear way - Asking for the answers to the evaluation questions and their evidence - Asking for an explanation of the limits of the findings and how general they are - Asking for policy recommendations in relation to the most reliable results
10.	**Producing the report**	- The report is too long, confusing with low readability - The report does not offer clear answers to the evaluation questions - The report does not allow the reader to form its point of view because of preemptive conclusions throughout the text - Results are not based on clear evidence - The report is internally incoherent (differ sections mutually contradict themselves).	- The report should not be longer than 100-120 pages - The report has to follow a pre-defined structure (outline) pre-agreed with the MA and an annex to describe in detail the methodology - The report must cover all the evaluation questions and the text must be easily understandable - An adequate visualization and summary of main results should be provided - Recommendations are clearly linked to the evidence and related to their complexity (administrative costs, timing, feasibility, etc.)
11.	**Presenting the report**	- Lack of a serious discussion of the results in the Monitoring Committee - Lack of scientific validation of and weak confidence in the results - Difficulty in involving political and decision-making levels in the discussion of the evaluation - Beneficiaries and stakeholders are not aware of the results and don't have the possibility to discuss them - Lack of accountability and transparency in the presentation of the results	- Providing different and adequate venues to discuss results (e.g., the National Evaluation Working group) not only the MC - Involving all the key stakeholders in the discussion - Organizing workshops (potentially with academia or with other experts) to validate results - Making the report accessible to laypeople through adequate syntheses and media
12.	**Using the findings**	- Evaluation findings are not discussed, even when the evaluation is important and reliable - Evaluation findings are misinterpreted and used in the wrong way	- Discussing recommendations with the implementers (the first users and as connection to the policy makers) - Asking for a specific and formal discussion of the recommendation in the National Evaluation Working Group and/or in the Monitoring Committee - Monitoring if recommendations approved by the MC are implemented in time and properly

Source: Developed by Authors

In conclusion, the section 1.5 of the evaluation plan on evaluation implementation and quality (see annex A) should include:

- A clear timetable and the organizational arrangements to respect it. As indicated in the template, a summary table of the planned evaluations must be provided including information on their timing and estimated duration; in addition, a GANTT can be annexed to the plan to offer a more detailed view of the sequence of evaluations and their main implementation activities.

- The main procedures and arrangements put in place to reduce the risks in the evaluation process and provide a high-quality evaluation. These elements can follow the issues and the possible mitigation measures suggested in Table 15, but adapted to the specific needs of the evaluation plan.

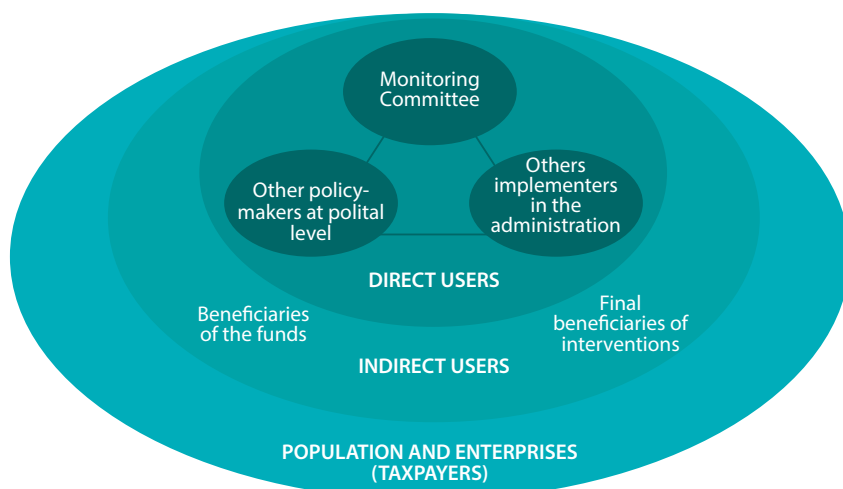## 5.4. DISSEMINATION AND USE OF THE FINDINGS

5.4.1. Communication of evaluation results

Dissemination of the evaluation results and involvement of the policy makers are key aspects for the success of an evaluation. EC regulations do require that evaluation reports must be published on the website of the OP; however, this may be insufficient to disseminate findings and recommendations. It is important to emphasize that access of key players to evaluation results and policy implications, is key to the success of evaluations.

Identifying the users of the evaluations from the start of the planning stage will set up the foundation to disseminate information; it allows time to involve the proper stakeholders, pose the right questions and organize an adequate strategy to communicate the findings. Users and stakeholders are not necessarily the same, even if they can largely overlap; in general, users must have the power to modify the programme, or the intervention, and this may occur in a direct way (political or administrative policy makers and implementers) or an indirect way (beneficiaries and social partners) and with a different influence. Democratic and participative principles inspiring the EU evaluation require to ensure the easy access of the population and enterprises (especially those living in the territory affected by the evaluated intervention) to the evaluation results. The main "institutional" user is the Monitoring Committee which includes the essential policymakers and stakeholders and has the obligation of knowing and discussing the findings of the evaluation; the Monitoring Committee has also the power to take decisions affecting the OP based on the suggestions taken from the evaluation. Figure 8 shows the main users of the evaluation according to their influence on the OP.

FIGURE 8 **The "constellation" of the users of ESIF evaluation**

Monitoring Committee

Other policy-makers at polital level

Others implementers in the administration

DIRECT USERS

Beneficiaries of the funds

Final beneficiaries of interventions

INDIRECT USERS

POPULATION AND ENTERPRISES (TAXPAYERS)

Source: Developed by Authors

Communication of the evaluation results needs to be carefully planned and adapted to different users. To be effective, it must respond to the information needs of the targeted users, be intelligible by the different audiences and useful to their interests and decision-making processes. This means that both content of the evaluation report and means of communication need to be adapted to the different audiences; moreover, the presentation of the results should match the timing of their needs (e.g., mid-term review of the OP, revision of measures to reinforce investment decisions of enterprises, preparation of the new programming period, etc.).

Ideally, communication is a two-way process where questions can be asked and answered from both sender and receiver, helping to further progress with the information gained. A direct involvement of the policy makers in the preparation of the evaluation when evaluations questions are collected and in the dissemination activities (such as, dedicated workshops or wider debate with the different stakeholders at central and regional level) can open decision-making processes to changes and improvements suggested by the evaluation results.

Different channels and venues can be used to reach all the different types of audience. The channels include:

- **Publication of the evaluation reports on the dedicated website -** According to the EU CPR, this is compulsory and allows all the interested citizens, organizations and academic experts to access the evaluation results. High readership requires that the evaluation reports were easy to understand but scientifically complete (with annexes including used data and methodological details), and able to provide evidence to sustain the judgments and policy recommendations. Visualization and other presentation devices should facilitate the comprehension of the reports to a broad audience. Availability of the evaluation reports can be advertized with the help of informative mail to the main potential users.

- **Adapted and targeted summaries –** Different policy makers may demand different syntheses of the evaluation results. For instance, politicians generally need to focus on a few principal messages and the consequent policy implications emerging from the evaluation; thus, a summary of two to three well-focused pages is sufficient. In contrast, implementers need to know if the implementation mechanisms must be revised or if the overall strategy must be improved and the two aspects may rely on different offices or administrations; in this case, more

articulated explanations of the results are necessary. Then, experts or academic scholars may be interested in replicating or extending the analyses carried out in the evaluation and need methodological features and data. Also, visualization options for evaluation results can be adapted to the targeted audience. In general, more specialized audience categories should be offered with more details or technical versions of the reports, whereas decision-makers, or less specialized audiences could be provided with more user-friendly options, such as info graphs and fact sheets highlighting not only the key data, but also some illustrative cases or empirical examples.

- **Organization of meetings and workshops with users/stakeholders -** These can be organized with the different categories of users and stakeholders, depending on their specific interest or involvement in the topic covered by the evaluation. These can include line ministries, decision-makers, representatives of the business environments (if the evaluation covers themes relevant for the sector), academia, etc. Materials for the meeting and presentation should be adapted to the specifics of each audience. Discussions and debates should be encouraged, to create room for increased awareness of the evaluation results and enhanced use of results. The objective of the workshops can change according to the audience (e.g., to influence policy design with Monitoring Committee and policy makers, to validate and reinforce results with experts and academics, to promote widespread participation of the policy with social partners, etc.) and consequently, the presentation and the discussion must be adapted to the profile of the participants.

- **Discussion of evaluation findings and recommendations in the Evaluation Steering Committee (part of the MC) –** The composition of the Evaluation Steering Committee, which generally includes the members of the MC on a voluntary basis and key invites (such as academia, senior evaluators, other relevant stakeholders for the topics evaluated) is expected to promote an extended debate and consultation on evaluation results. It will act as a forum for discussion in a context of multi-faceted expertise and prepare the more formal discussion in the Monitoring Committee. In particular, policy recommendations of the evaluation can be analyzed in this group to verify pros and cons and understand their feasibility; this allows submitting already screened recommendations to the Monitoring Committee.

Table 16 shows some of the possible activities that will help disseminate results among the different types of users. These suggestions imply an adaptation of the communication tools to the different contexts, although not all the proposed activities have necessarily to be implemented for all the evaluations; the tools have to be readjusted based on the quality and the reliability of the evaluation, its importance in the OP strategy and the more or less extensive involvement of local actors of the evaluated intervention(s).

TABLE 16 Possible activities to disseminate evaluation results by type of users

Type of user	Possible activities
DIRECT USERS	
Monitoring Committee	- Communication and presentation of results in the meetings of the MC
	- Preparatory and more in-depth discussion of the evaluation results in the national Evaluation Working Group
Other policy makers at political level	- Targeted "policy brief" documents on the main messages emerging from the evaluation
	- Restricted seminars and meetings to explore findings and policy implications more in details
Other policy makers at administrative level (implementers)	- Targeted operational documents on results and recommendations proposed in the evaluation report
	- Workshops with key administration representative to discuss findings and policy implications
	- Presentation of the results with the EC or at EU level to verify and promote their generalisation ("peer reviews")
INDIRECT USERS	
Beneficiaries	- Annual conferences with social partners on the evaluation results and other relevant studies
	- Seminars with organizations of beneficiaries and social partners
	- Meetings or focus groups with beneficiaries involved in the evaluations, as feedback of their contribution
Final beneficiaries	- Meetings or focus groups with final beneficiaries involved in the evaluations, as feedback of their contribution
	- Meetings with territorial/sectoral organizations and municipalities to improve interventions and awareness of the evaluation results
SOCIETY ("TAXPAYERS")	- Public access to the report and (user friendly) summaries
	- Press conference and dissemination of results through media (when useful)

Source: Developed by Authors

5.4.2. Use of evaluation findings and recommendations

Systematically monitoring how evaluation findings are used for decision-making could also be helpful in improving the evaluation process.

A first aspect to be considered relates to the selection of recommendations that need to be implemented. This should be done in the Evaluation Working Group, through a detailed review of the evaluation reports and the formulated recommendations. At the OP level, the Evaluation Steering Committee (which could include the members of the MC) could play a key role to this end. Organization of a workshop with the implementers to discuss the main recommendations selected for implementation might also be beneficial.

To monitor implementation, MAs and other relevant stakeholders could be asked to complete a standard and short monitoring form, containing information on the extent to which evaluation findings/recommendations were used and the reasons behind the lack of use (e.g., poor quality evaluation, not feasible, etc.). Progress must be monitored annually, in a centralized manner, by each MA and at ECU level.

The improvement in the use of the evaluation results needs the involvement of implementers (administrative representative responsible for the implementation of the evaluated interventions, which is not necessarily the MA) and of the policy makers at the political level. In this respect, the above monitoring should include a note on the meetings and the presentations organized by the MA to present the results of the evaluations to these key actors and the consequent decision taken, if any.

Here mentioned are arrangements aimed to inform the section 1.6 of the evaluation plan on dissemination and use of the evaluation findings. They should be adapted to the OP and the planned evaluations and their financial and administrative efforts must be compatible with the capacity and objectives of the MA. The dissemination activities are an essential part of the plan and must be well integrated with the governance of the plan (MC and Evaluation Steering Group) and with the entire evaluation process involving users and stakeholders from the evaluation plan.

5.5. REINFORCING EVALUATION CAPACITY

Developing evaluation capacity must be a shared concern of the wider policy community. Activities in this area should thus cover both those who manage and commission evaluations, as well as actors with an interest in evaluation results at a policy and programme level and entities who undertake evaluations.

Evaluation capacity is multi-faceted, and actions need to be implemented at many different levels that reinforce each other:

- *At the individual level, by ensuring the necessary skills and competencies* (through activities aiming at skills development such as training courses, coaching and other activities of this type)

- *At the organizational level, by ensuring that adequate management arrangements and structures are in place;* Evaluation should be integrated at each stage of policy and programming: from planning through to implementation and follow-up. Evaluation findings should support decision-making, in terms of deciding what policy options to choose, how best to implement and when assessing what has been effective.

- *At the inter-organizational level, by bridging public and private bodies through networks, procedures, and partnerships.* Coordination should bring consistency in the way evaluations are commissioned, managed, and executed across government and ultimately across the public sector. Other key aspects refer to the existence of a well-defined market with clear rules, as well a culture of evaluation that values professional standards, independence, learning from experience and evidence-based policy.

- *At the societal level, by embedding evaluative thinking in civil society (including professional organizations) and in the public sector.* That way, open and systematic dialogue must be maintained between policy makers and evaluation specialists, so that priorities for evaluation can be identified and scrutinized. The existence of an evaluation community of practice and evaluation associations are key to reinforcing a culture of evaluation and for an increased awareness of evaluation activity and outputs to support an enhanced use of results.

Actions to develop evaluation capacity must follow an integrated approach, covering four pillars:

- **Institutional set-up:** this will be done by adequately locating and structuring evaluation functions and ensuring their coordination.

- **Strengthening evaluation demand:** ensuring that there is an effective and well-managed demand for evaluations (from both government and civil society)

- **Strengthening evaluation supply:** ensuring that the skills and competencies are in place with appropriate organizational support (from educational institutions, but also by strengthening the community of consultants and supporting the development of a professional evaluation community). High standards need to be encouraged throughout the process and the prerequisites to achieve this must be ensured (evaluation culture, independence of evaluation involvement of academia, professional associations, development, and application of high standards, etc.).

- **Institutionalizing evaluations:** building in evaluation to policy making systems and across the broader policy system. This requires on one hand, to extend the use of evaluation within the public sector and, on the other hand, to better integrate evaluation processes and use of evaluation results into policy making and programme management.

A well-developed evaluation system entail:

- A strong commitment to learning lessons and improvement.
- Avoidance of a blame-culture which discourages learning.
- The commitment of the policy makers to evidence-based polices in the broadest sense.

- A commitment to excellence, high standards and continuous improvement.
- Evaluation used to promote transparency and multiple accountabilities to the public and state and regional authorities.
- A commitment to learning from both the government and public sector.

Capacity building involves actors inside and outside the administration, as well as contracting authorities (MAs, ECU-MEIP) and evaluators. It is expected that a significant part of capacity building activities will be carried out by the ECU at national level, involving all the decentralized MAs and administrations. Nevertheless, important activities should be carried out also at MA level. Table 17 presents the activities that could be implemented for each of the pillars described above.

TABLE 17 Activities for reinforcing capacity (at national and OP level)

At national level (by ECU)	At OP level (by the MA)
Institutional set-up	
- Training and capacity building activities for the staff with evaluation responsibilities (in ECU and MAs) - Providing for a well-functioning coordination mechanism (guidance, methodologies, regular meetings to discuss challenges and best practices, meetings to discuss progress and next steps arrangements for common/complementary evaluations, in case the need is identified,)	- Defining key collaboration procedures between the valuation unit and the other departments within the MA - Ensure participation of key staff in the evaluation units in the actions aiming at capacity development and to the events organized in the framework of the national evaluation working group.
Institutional set-up	
- Ensure all necessary arrangements for the adequate functioning of the national Evaluation Working Group (procedures, resources, etc.) - Ensure resources are available for the functioning of the Scientific Committees (at the level of each evaluation or only for complex evaluations)	- Provide for general training of the MA and IB staff on general evaluation topics, as to foster increased use of evaluation throughout the programme cycle, including in planning and implementation and generally, in policy making - Ensure resources are available for the functioning of the Scientific Committees (at the level of each evaluation or only for complex evaluations)
Strengthening evaluation demand	
- Actions to increase understanding of the evaluation use in policy making, especially among decision-makers - Ensure a better alignment of the evaluations with the information needs and make sure results are provided in time to support decision-making - Ensure resources for evaluation are available	

| At national level (by ECU)	At OP level (by the MA)
Strengthening evaluation supply	
- Promote evaluation quality standards, that must be followed by the all the actors involved in the evaluation process - Continue to support the Romanian Evaluation Network, as an opportunity to bring together professionals for the evaluation commissioners, academia, and evaluators. The Network also acts as a forum for discussion of best practices and challenges in evaluation. - Meetings and networking to exchange experiences at national and international levels (incl. on methods on instruments, best practices and lessons learned, etc.) - Connection and exchanges with the Romanian Evaluation Network and other key stakeholders (public authorities, evaluators, academia etc.) to promote and share evaluation culture.	- Meetings and networking to exchange experiences at regional or sector level (including on methods on instruments, best practices and lessons learned etc.) - Connection and exchanges with the Romanian Evaluation Network and other key stakeholders, including at local and regional level (public authorities, evaluators, academia etc.) to promote and disseminate evaluation culture.
Institutionalizing evaluations	
- Organization of regular meetings and networking events with different categories of stakeholders (social partners, other administrations, academia, evaluators etc.) to increase awareness of evaluation, discuss evaluation findings and the need for future evaluations - Systematic follow-up of evaluation findings and recommendations such that middle managers pay more attention to the evaluation results and justify why they may not be following recommendations - The extension of evaluation throughout the policy cycle from option identification to planning, programming, implementation and reprogramming the design of the next phase of policy.	

Source: Developed by Authors

The strategy to strengthen evaluation capacity must be presented in section 1.7 of the evaluation plan. It includes the main gaps that need to be filled in, the different actors to support, the initiatives to promote and the intended results. This strategy at OP level should to be coordinated with the national strategy of ECU and, consequently, clearly delineate its scope and purposes to the capacity relevant for the OP and the related sectors/territories involved.
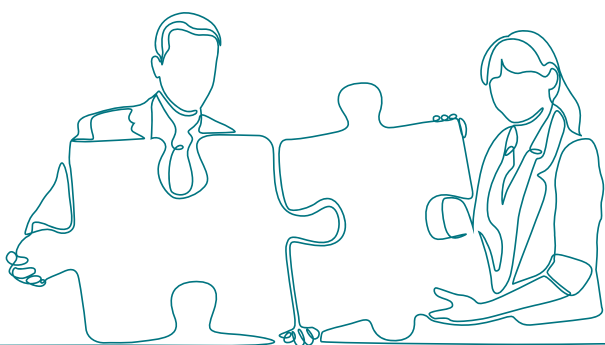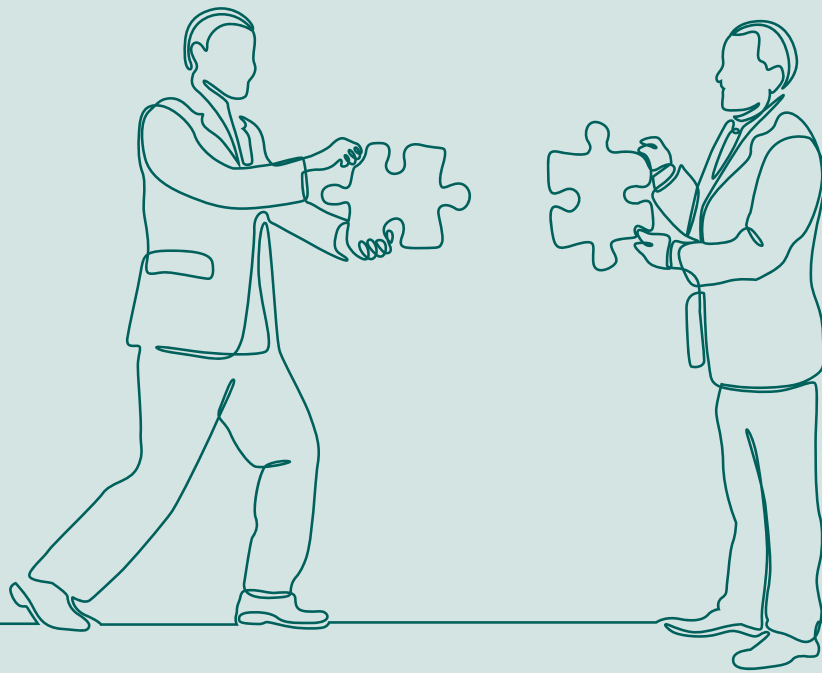
HOW TO IMPLEMENT AN EVALUATION PLAN

6

The implementation of the evaluation plan requires several different activities. The main activities (defining the evaluation questions, identifying the type of evaluator, and selecting the evaluator, ensuring the evaluability of the interventions, supporting the collection of information, supervising evaluators and reviewing evaluation reports, using the evaluation findings) have been briefly described in the section 5.3 on the evaluation quality. The current section describes two key activities in detail:

i.   the selection of the external evaluators including the terms of reference for the calls for tender;

ii. the basic structure of an evaluation report.

6.1. THE SELECTION OF EXTERNAL EVALUATORS

The process of selecting external evaluators needs to be transparent and objective. Careful drafting of the Terms of Reference (ToR) is a key factor to attain this goal, together with the use of a formal selection committee. This should include representatives of the people in charge of the evaluation and, when possible, representatives of the potential and actual users of the evaluation. Sometimes, it is useful to also include a technical expert if the formal committee lacks sufficient expertise.

The selection of an external evaluator requires three main steps:

1. Identifying the most suitable procedure for selecting the evaluator.

2. Preparing the public procurement documentation, including the Terms of Reference of the call for proposal.

3. Assessing proposals and selecting the evaluators.

The following paragraphs present the main activities needed for carrying out an effective selection process. The paragraphs are oriented towards the aim to select the best quality proposal. They aim to convey objective and measurable elements to assess the bidders' know-how and operational capacity.

6.1.1. Identifying the most suitable procedure

The identification of the public procurement procedure[19] responds to several exigences and must balance some trade-offs:

- a *negotiation without prior public notice*[20] is generally quicker than an open one, but by nature restricts competition and can be used only for a limited amount of expenditure.

- an *open procedure*[21], on the other hand, can be used for large evaluations or multiple-evaluation contracts, but it is relatively more complex and time consuming.

- a *framework agreement*[22], which is normally assigned following an open procedure has a duration of up to four years, which could cover most of the programming period and cover more than one evaluation. This procedure allows for the selection of one or more bidders[23] who will compete in restricted and fast procedures along the duration of the agreement; in this way, agreements limit competition but facilitate selection processes.

In order to receive good quality proposals, it is generally advisable to make use of open procedures, which increase the number of candidates. Nevertheless, these procedures can be limited in the case of small studies or very specialized evaluations where there are only a few qualified candidates. Clearly, legal constraints play an important role in deciding the procedures and the verification of their full respect is needed.

The Romanian legislation requires that the proposals are evaluated according to the "price-quality ratio" criterion[24]. Consequently, the identification of the public procurement procedure also requests the definition of the "price-quality ratio" to use. This ratio is determined by the score assigned to the price criterion and the score of the technical criterion, which measures the quality of the proposal; in general, the price criterion cannot exceed 40 points on 100, the remaining score being assigned to the technical criterion (minimum 60%). These two criteria have the following characteristics:

---

19 On these aspects see also: European Commission, DG Regio, 2015, Public Procurement - Guidance for practitioners on the avoidance of the most common errors in projects funded by the European Structural and Investment Funds; https://ec.europa.eu/regional_policy/sources/docgener/informat/2014/guidance_public_proc_en.pdf

20 See para. 7, art. 104 of Law 98/2016, available at: http://anap.gov.ro/web/legea-nr-982016-privind-achizitiile-publice/;

21 In open procedures, any interested economic operator may submit a tender in response to a call for competition. The minimum time limit for the receipt of tenders lags between 10 days (for simplified procedure) and 35 days (for open tender), from the date on which the contract notice was sent.

22 A framework agreement means an agreement between one or more contracting authorities and one or more economic operators, the purpose of which is to establish the terms governing contracts to be awarded during a given period, in particular regarding price and, where appropriate, the quantity envisaged. The term of a framework agreement shall not exceed four years, save in exceptional cases duly justified, by the subject of the framework agreement.

23 From only one to many proposals, depending upon the number of qualified and accepted proposals.

24 As stated in the recital of the EU Directive n.24/2014 on public procurement: "In order to encourage a greater quality orientation of public procurement, Member States should be permitted to prohibit or restrict use of price only or cost only to assess the most economically advantageous tender where they deem this appropriate."

- The *price criterion*. This is generally based on the total budget of the proposal, but a more detailed analysis of the budget (for instance, by task or by report) could also be considered. The mathematical formula used for assessing the budget is also critical, since depending on the algorithms, price differences of candidates may be emphasized or minimized. In general, it is necessary not to exaggerate the weight of the price component (20% may be a right weight) because it would reduce competition focused on quality and could result in low quality evaluations. This criterion should play a limited role and help to select similar and high-quality proposals.

- The *technical criterion*. This should be used especially in the context of the intellectual services, as evaluation services are. According to the Romanian legislation, the price criterion alone cannot be used for intellectual services with a high degree of complexity[25]. The technical criterion should comprise non-discriminatory, objective, and quantifiable evaluation factors. The technical criterion sums up the scores of several detailed technical criteria such as the quality of the methodology, the experience of the team, and the proposed quality assurance system (see below).

6.1.2. The Public Procurement (PP) Documentation

The Public procurement documentation has a two-fold objective:

a. Making the request of services clear and complete to receive high quality and comprehensive proposals.

b. Facilitating the selection of the candidates with comparable and objectively judgable requests.

According to the Romanian legislation (i.e., Law 98/2016), two main types of documents are used in the PP process, each with its own dedicated function, scope, and content, as it is presented below:

Data Fiche[26] containing mainly administrative specifications, such as:

i. How the procedure works;

ii. Deadline and means foreseen for sending the proposal;

iii. The compulsory eligibility (qualification) criteria (both financial and technical), quality references to be met

iv. The award criteria;

v. The content of the documentation to submit, which details on the structure and the length (in pages and when useful) of the technical proposal to submit, the way to submit and receive answers to clarifications, as well as deadlines for questions and/or the schedule for a bidders' conference.

Technical Specifications or Terms of Reference[27] (ToR), containing specific requirements and elements to be considered when drafting the technical proposal. The latter should be structured according to the main elements of the ToR.

The Data Fiche presents eligibility and quality criteria, as well as evaluation factors, which are then detailed in ToR. The Data Fiche briefly presents all the elements needed for a bidder to submit an offer while the details are further provided in the ToR.

- The *eligibility (qualification) criteria (or Data Fiche)* used to assess the potential participants; these criteria generally refer to the experience and the operational capacity of the bidders to provide the requested service. They are measured by indicators like the number of similar services provided and/or the turnover of the candidates.

  These criteria must be defined according to the existing legal framework, but their definition includes important discretionary elements (e.g., what is a similar service or what is the most appropriate level of turnover to judge the capacity of a tenderer). To ensure a wide participation and a real competition, it is important that the definition of these criteria is not too strict and is proportional to the complexity of the evaluation. The eligibility and/or qualification criteria included in this section must not be doubled and superposed with similar ones included in the quality or selection criteria.

The *quality criteria (technical requirements or Terms of Reference),* to select the proposals; these criteria are focused on the specific characteristics of the evaluation services and in particular:

a. The *context of the evaluation* – to assess the knowledge of the candidates on the main objectives and the policy context of the requested evaluation, together

---

25   As per art. 187, para. (8) of Law 98/2016
26   As per Romanian legislation, depending on the type of tender, this section is translated as *"Fisa de date a achizitiei";*
27   As per Romanian legislation, depending on the type of tender, this section is translated as *"Caiet de sarcini";*

with their understanding of how the intervention works and the awareness of the socio-economic factors can potentially influence the outcomes of the intervention.

b. The *proposed methodology* – to assess the quality of the methodological and operational solutions proposed by the candidates in relation to the methodological approach, the data to use, the tasks to fulfill, the reporting and the timing of the evaluation.

c. The *proposed team* – to assess the quantity and the quality of the human resources mobilized for the evaluation. In this section, it is important to consider the following: a) the qualifications of the experts proposed, based on their CVs and credentials, b) the balance between senior and junior consultants; 3) the mix of different experiences; 4) the planned efforts of the different consultants.

d. The *quality assurance arrangements* – to assess the capacity of the candidates in facing the challenges raised by the evaluation. The quality system of the evaluator should be judged based on its capacity to identify the most relevant risks of the evaluation (e.g., delays in data availability, possible changes in the team, reluctancy of the survey respondents in giving information, necessity for new analyses, etc.) and suggesting reasonable and feasible mitigations or solutions to risks.

The Contracting Authority will select the best offer out of the admitted offers by applying the evaluation factors, which is detailed in ToR. The Data Fiche must specify the award procedure, which in general consists of three main steps:

- Verification of the compulsory eligibility criteria (as detailed in the Data Fiche). Examples of such criteria include the following: no outstanding debts to the state budget, a minimum level of turnover for the last one to three years, etc.

- Verification of the compliance with the minimum level of quality criteria (detailed in ToR). For example, similar experience requires a minimum one to three finalized contracts, involvement of at least three technical experts with a minimum of three years' experience in evaluation, etc.

- Scoring the offers based on the evaluation factors. These can be based on the quality criteria, or other elements. For example, scoring the technical experts' expertise in the evaluation field: three points for three years of experience, five points for three to five years, 10 points for over five years of experience, etc.

In general, the Terms of Reference (technical specifications) include[28]:

Background information and rationale - The opening section of the technical specification in ToR typically provides an orientation about the overall programme, project, or activity to be evaluated. Depending on the complexity of this programme and the context for this evaluation, this section might be a few paragraphs or a couple of pages.

Details should focus on the following as appropriate:

- The current objectives and intended outcomes of the programme being evaluated.

- The rationale for the evaluation and the key overarching evaluation objective and questions, including an overview of what decisions might likely be influenced by the findings.

- A history of the programme, including how these objectives and targeted outcomes have changed over time.

- The context in which the programme is situated—including organizational, social, political, regulatory, economic or other factors that have been directly relevant to the programme's implementation.

- The roles and responsibilities of various key stakeholders in designing and implementing the programme, noting any significant changes that have occurred in these roles over time.

- Any studies or evaluations that have been conducted on the programme or related activities to date. If available, the monitoring and evaluation framework for the programme should be attached.

Objectives of the evaluation and evaluation questions - The framework and presentation of the objectives of the evaluation usually represent a brief, but important section in any ToR. A clear, concise formulation of both the objective and the evaluation questions are critical to facilitate the preparation of relevant technical offers. A common understanding of, and consensus around, the stated objectives and evaluation questions

---

will be important throughout the implementation of the assigned tasks. Objectives for evaluating OPs/PAs must relate to aims (learning/accountability/knowledge), needs of the users and regulation requirements.

The statement of the specific objectives for the evaluation should adhere to the following guidelines where possible:

- *Avoid a lengthy list of objectives.* A simple mission or evaluation of an activity might focus on a single objective. In broader studies, the focus should be limited to no more than three to five objectives;

- *Use clear* outcome-*focused language.* The objectives should not be stated in technical or process terms. Specific evaluation questions should be identified for each objective presented by the ToR. Depending on the nature of the evaluation, these questions could range from being broad to being quite specific (Broad question—What has the program achieved vis-à-vis its objectives? Specific question—What is the impact of the program on minority girls' enrollment and dropout rates?)

To the extent possible, the list of questions should be kept to a minimum; this will allow for room when unexpected issues will arise and the evaluator needs to refine the list and specify possible additional questions in exploring the purpose, scope, and methodology. There should be a logical progression between the purpose of the evaluation, its specific objectives, and the questions posed in relation to each objective. Questions should be specific and focused on the activity being evaluated. In constructing the list of evaluation questions, it is important to prioritize these areas of inquiry according to the information needs of stakeholders and the overall rationale driving the evaluation.

Scope of the Evaluation - This section presents the parameters of the evaluation in terms of its scope and limits. The scope should be realistic given the time and resources available for implementing the study. Details here could include:

- *Time period, territories and programme components* covered by the evaluation.

- *Other existing or planned evaluations* of the same subject. In the interest of uniformity, this study could build on or complement related activities being conducted by development partners.

- *Target groups.* The evaluation might focus on a subset of beneficiaries or on the complete range of engaged stakeholders.

- *Issues that are outside of the scope.* The commissioner of the evaluation might recognize some aspects of the program or local context that will be difficult or impossible to analyze during the prescribed time period. The ToR should clearly specify when such topics are not a focus of the intended tasks.

Approach and Methodology - Specifying the approach for the evaluation can be the most challenging part of developing the ToR. This section should outline how the evaluation will be conducted. However, many ToRs leave room for the evaluator(s) to define a more detailed methodology in line with the prescribed scope and objectives. Key elements generally highlighted here include:

- *The overarching methodological framework* (for example, case study, sample survey, mixed methods, counterfactual impact evaluation, theory-based evaluation and so forth).

- *Expected data collection and main analysis methods* with descriptions of any instruments used to collect needed information.

- *Outcome and output indicators* that are being proposed or have been used to measure performance, along with associated baseline and target data

- *Availability of other relevant data*, such as existing local, regional, or national data, or data from similar programs

- *The process for verifying findings* with key stakeholders

- *How various users/stakeholders/reviewers are likely to be involved in the evaluation* (for example, steering committee, scientific committee).

- *Dissemination* of results

The expected approach should be described with enough flexibility to draw on the expertise of selected evaluators. For a competitive process, those responding to a ToR can be asked to submit a more detailed proposal for the evaluation plan. This process during the selection phase will allow candidates to differentiate themselves in terms of the quality of their proposed methodologies. The ToR should also ask the evaluator(s) for a detailed methodological approach as the first deliverable of the evaluation.

The methodological section can also specify the tasks that the evaluator must fulfill to complete the evaluations. Tasks are defined according to the steps needed to implement the methodology and deliver the expected reports and outputs.

Professional Qualifications - The mix of knowledge, skills, and experience needed will depend on the scope and methodology of the evaluation. The ToR should specify as clearly as possible the desired profile of the evaluator or team to attract the strongest candidates for conducting the study.

The requirements listed for the expert candidates should directly be drawn from the skills required to carry out the activities as well as factor in the complex and specific nature of the evaluation(s) subject.

Relevant and useful details in this section relate to:

- *Number of experts needed to successfully implement the required tasks (whether a limited number of experts or a larger team is needed).*

- *Qualifications needed: What specific expertise, skills, and prior experience* is the evaluator (individual or team) required or desired to possess — including evaluation skills (qualitative and/or quantitative methods), technical competence in the sector to be evaluated,

process management experience, language proficiency, and in-country or regional experience

- Team interaction: *how the different expertise, skills, and experience among team members* will be integrated and complement each other

- *Distribution of responsibilities: What the expected distribution of responsibilities* is among the team leader and team members.

Any additional information that will help gauge the qualifications of potential evaluators should also be noted in this section. This includes requests for curriculum vitae, references, or examples of evaluation reports recently completed.

Within the framework of the Romanian legislation, the National Authority for Public Procurement (ANAP) has issued clarifications[29] with respect to the technical criteria, especially qualifications, expertise, and related underlying documents for experts.
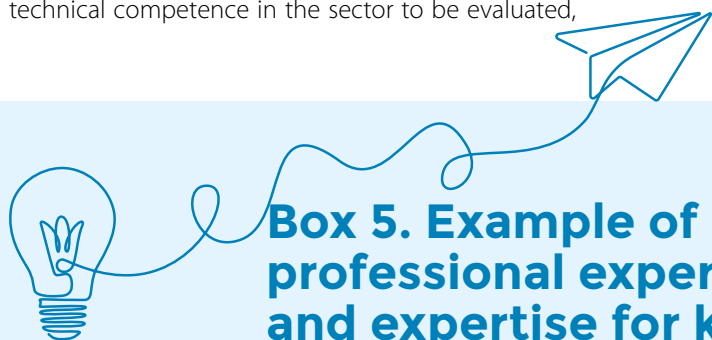
Box 5. Example of requirements for professional experience, qualification, and expertise for key experts

Requirements for professional experience, qualifications & expertise for key experts (example)

Key Expert #2: Team manager

- General professional experience in consultancy services: at least five years.
- Specific professional experience (expertise) in project management: at least three years.
- Professional qualifications: certification in project management (issued by the Project Management Institute (PMI) or similar).
- Similar experience: involvement in a similar position in at least three contracts requiring project management related activities.
- At least three verifiable recommendations confirming professional and similar experience.
- Key Expert #2: Technical evaluator (e.g., for construction contracts in investment-related projects)
- General professional experience: at least five years.
- Specific professional experience (expertise) in the field of evaluation of investment-intensive projects: at least three years.
- Professional qualifications: Bachelor of Arts in constructions (civil engineering).
- Similar experience: involvement in a similar position in at least three contracts requiring evaluation of investment-intensive projects.
- At least three verifiable recommendations confirming professional and similar experience.

---

29   Instruction 1/2017 issued by ANAP, available at http://anap.gov.ro/web/wp-content/uploads/2017/01/Instructiunea-nr-1_2017.pdf

Detailed requirements could relate to both general and specific experience; also, task-related professional experience and similar experience should be requested, especially for a limited number of team members (usually referred to as "key experts", such as, for instance, statisticians or evaluation methodology experts).

Deliverables and Schedule - The outputs and reporting requirements expected for the evaluation should be specified, along with the required or proposed timeline for the study. Clear guidance in this section will help ensure that the outputs from the evaluation meet expectations. Details should include the following:

- *Specific information about the types of products* (reports, presentations, and so forth), the expected users, and how the different products will be used. Ideally, each product will be listed separately to specify its individual requirements and timeline.

- *The structure and format* for each product. This would include any expectations regarding length and content (for example, the order of sections or the inclusion of an executive summary).

- *The language(s) in which deliverables should be written* (for instance executive summary may be requested also in English for a major dissemination)**.**

  - *Organizational standards and practices.* Any established style guide or standard formats for written documents should be referred to here, including code of ethics or established principles directly relevant to the evaluation (for example, transparency, confidentiality, and so forth).

  - *The timeframe for products, including milestones.* ToRs must specify when a project is expected to launch, when a detailed evaluation plan or inception report will be expected, and so on. The estimated due dates should be clearly specified for each activity and product. Alternatively, some ToRs provide an overview of the expected scope, objectives, and deliverables and request that evaluators propose a realistic timeframe.

  - *Required meetings/consultations.* This section should also note the frequency and types of meetings expected with the evaluation commissioner and other stakeholders of the evaluation.

Regardless of which deliverables are requested, the description of the requested outputs and prescribed timeline should allow for ample opportunities and time for peer reviews and other feedback from stakeholders to be incorporated into product revisions.

Budget and Payment - The commissioner must specify the maximum cost payable for the evaluation and the conditions to define the budget proposal (the admissible costs, the presentation of a total cost or costs of the main invoices) and the terms of payments. These elements will be considered to assess the price-quality ratio of the proposals according to a predetermined mathematical formula30.

- In cases where a limited budget will likely constrain the scope and methodology of the study, an effective practice is to state the available budget and ask proposers to describe what they can expect to achieve. Such an approach allows for a price-quality ratio assessment.

- Aside from information on the budget itself, this section includes any pertinent details related to payment. For example, the type of contract (for example, fixed price or time plus materials), the required process for invoicing, or the intervals and means of payment might be specified.

Structure of the Proposal and Submission Guidelines
ToRs must be used to request proposals from potential evaluators as part of a competitive process. Consequently, the ToR should provide instructions regarding the proposal format, content, and submission process, including details on the Structure of the proposal, including the list and order of topics to be covered. The structure should follow the selection criteria to facilitate the work of the selection committee; for instance, it may include:

- The understanding of the service (objectives, scope, rationale, etc.)

- The socio-economic and policy context of the evaluation

- The proposed methodology (the methodological approach, the information to collect, the tasks to fulfill, the specific techniques to apply in each task and the expected results)

- GANTT summarizing the activities

- Schedule of the deliverables and their planned structure

---

30 In cases where a limited budget constrains the scope and methodology of the study, an effective practice is to state the available budget and ask proposers to describe what they can expect to achieve. This "inverse" procedure allows the same for price-quality ratio assessments.

- The proposed team (components and CVs, organization and roles in the team, interactions with the commissioner, planned effort)

- The quality assurance arrangements (quality arrangements, identified risks, mitigation measures)

- The financial offer (total or by tasks)

6.1.3. Assessing proposals and selecting the evaluators

Table 18 below presents an example of how the different evaluation factors can be weighed to assess the offers submitted by the bidders. It proposes a range of different weights that can be assigned to each criterion; the combination of the points in the different criteria can change significantly according to the type of evaluation and the selection procedure defined. The weights assigned to all criteria have to make up a total of 100 points. However, the priority of the methodological quality and the significance of the other elements should always be safeguarded to promote high quality evaluations.

Criteria should be clear, concise, objective, comparable and measurable. Above all, they must not favor any potential candidate by proposing criteria which could not be fulfilled or attainable by a large majority of the potential tenderers.

Any criteria set-out as an evaluation factor must not be duplicated and used also as an eligibility (qualification) criterion[31]. For instance, similar professional experience of one tendered should be an eligibility criterion, while the specific experience and expertise of the assigned experts in the team should be a selection criterion.

TABLE 18 Possible range of points to assign to different criteria

Criteria and possible specifications	Possible points
The understanding of the evaluation context	10-20
- Quality of the analysis related to the assignment, in particular its relevance and coherence	
- Relevance of the tender and its expected results to the objectives of the requested evaluation	
- Knowledge of the intervention to evaluate and the administrative and socio-economic context in which it is implemented	
- Understanding the aims of the evaluation and of the expected use of its findings	
The proposed methodology	40-60
- Appropriateness, feasibility, adequacy, soundness of the proposed approach to respond to specific assignments requested	
- Quality of the methodology proposed to respond to the evaluation questions	
- Pertinence and quality of data proposed for reaching the evaluation conclusions evidence and carry out needed analyses	
- Clearness and quality of the synergies between the tasks to undertake	
- Efficiency and effectiveness of data collection and interpretation methods	
- Quality of the methodology and deliverables proposed as a support to the conclusions reached and their appropriateness, feasibility, and robustness.	
- Overall coherence between the work plan, methodology, allocation of time, resources, and quality control measures	

---

31   This is an imperative restriction set-out in Government Ordinance 395/2016 – Norms of the application of the public procurement Law 98/2016, Para. 7, art. 32

Criteria and possible specifications	Possible points
The team and the work organization	20-30
- Experience and expertise of the proposed team (for each key team member and/or for the team as a whole).	
- Specific professional qualifications to carry out the evaluation;	
- Adequacy and distribution of roles and responsibilities of the proposed team and of the different economic operators (in case of joint tenders, including subcontractors) for each task.	
- Global allocation of time and resources to the project and to each task or deliverable (The tender should provide details on the allocation of time and human resources and the rationale behind the choice of this allocation).	
- Mechanisms for continuous service and coordination inside the team and demonstrated ability to complete the work well within the time available	
- Clearly defined working method and capacity to liaise and cooperate with all the relevant stakeholders	
The quality assurance arrangements	10-20
- Pertinence, completeness, and reliability of the quality control system applied to the service; in particular, tenderers will be assessed on:	
• The approach designed to ensure risk management	
• The approach designed to report on the implementation and maintain stable communication with the commissioner.	
• The approach designed to minimize risks in collecting, storing and analyzing data	
• The approach designed to the management of complaints and corrective actions	
- Relevance of specifications concerning the quality of the deliverables, the reporting quality check, and continuity of the service in case of absence of the member of the team.	
- Methods to ensure the validation of the information collected and the final findings and conclusions.	
The price[32]	20-30

Source: Developed by Authors

6.1.3.1. Assessment of the quality and appropriateness of the proposed methodology

The adequacy of the proposed approach and methodology to address and fulfill the requirements set-out in the ToR should be central to the appraisal and selection of the best candidate (tenderer). The selection committee can ensure this by checking each of the offers submitted by the tenderers against the points below:

• Were the respective assignments requested addressed in a clear manner?

• Was the methodology proposed sound, appropriate, feasible and adequate and in line with the importance of the questions raised in the ToR?

• Are all the evaluation questions addressed in a satisfactory and verifiable manner?

• Does the proposal include the adequate methods for the collection and interpretation of relevant and sufficient information?

• Is the methodology based on rigorous analytical techniques?

• Is the methodology consistent with the evaluation questions and the interventions under examination?

• Is the methodology able to address the evaluation criteria in an impartial manner?

Assessment on the quality and adequacy (appropriateness) of the method proposed need to be carried out by experienced personnel. An example of scoring and specifications of the different criteria have

---

32   In such tenders having as main object intellectual outputs, the weight of the price factor in the overall evaluation factors must not exceed 40%, as set-out in Government Ordinance 395/2016, Para. 7, art. 32.

been proposed above in Table 18; the specifications must be modified and adjusted according to the evaluation service requested[33].

To avoid the risk of being obliged to select poor quality proposals in accordance with current national public procurement legislation, the scoring of the methodology should allow for objective judgment and differentiation among the tenders submitted, to avoid that all or too many tenders score the maximum points available.

There are different tools to facilitate this judgment. For instance, the "evaluation matrix" is a tool that the evaluators

frequently use to summarize their methodological approach and it may be expressly demanded in the technical proposal. Table 19 shows an example of this matrix, but it can be specified in different ways according to the importance of the different elements. In all cases, it is important that it starts from the evaluation questions and detail the other elements according to them. In this way, it would also be possible to assess the quality of the proposal for each evaluation question and subsequently aggregate the individual scores.

TABLE 19 Example of matrix to summarize the evaluation methodology

Evaluation questions	Main indicators to answer	Needed Information	Sources of information	Method of analysis

Source: Developed by Authors

6.1.3.2. Assessment of key qualifications, experience, and expertise of the team

Key qualifications, similar previous experience, and specific in-depth expertise of the team of experts proposed by each tenderer are important, especially if the methods proposed are innovative or particularly complex. The capabilities of the team must be matched with the methodology proposed to avoid problems occurring while the evaluation is being implemented. However, there is a risk that exhaustive and excessive requirements would discriminate against new entrants and therefore, disregard the principles of equality and ensuring free access for all tenderers (by setting no entry barriers). In order to cover such risks, the requirements regarding key qualifications, experience and expertise should be balanced and equitable; this aim can be attained by avoiding rare, disproportionate qualifications (especially in the case they are not directly related to the

evaluations services), scoring a decent level of previous experience (usually, 5-10 years general experience and 3-5 years specific experience should be considered reasonable), requesting expertise in the areas/ domains of the calls of proposals but also allowing for equivalence and so on.

It is always good to pay attention not only to the presence in the team of highly qualified personnel, but also to the time that they are prepared to devote to the task. As evaluations are time consuming, the most qualified people will not undertake all the workload themselves. Time allocated by those with experience needs to be sufficient to provide supervision for those in more junior positions. Evidence of the proposed team having worked together successfully is also relevant.

6.1.3.3. Assessment of the price

Assessing the proposed price for the services is an important aspect of the selection process but should not be overestimated. The weight of the price (financial)

factor in the overall criteria must not exceed the weight of the quality criteria and must not exceed 40% of the total.

---

33   For instance, the criterion "Are all the evaluation questions addressed in a satisfactory and verifiable manner?" could adopt the following evaluation factors:

5= all the questions addressed in a comprehensive and coherent way, new and detailed questions are proposed. information sources well defined and example provided.

4= all the questions addressed in a comprehensive and coherent way, information sources well defined.

3= most of the questions addressed in a comprehensive and coherent way, information sources sufficiently defined.

2= some of the questions addressed in insufficient way, information sources not well defined.

1= few questions addressed in an insufficient way, information sources not completely defined;

0= few questions badly addressed; information sources not defined.

A second point worth noting is that the assessment should not only consider the total price, but also the different components of the budget. For instance, you may realize that the spending in one component may be too low and highly unrealistic, or too high for the activities involved. One may also look at aspects such as the unit cost of the workday for the different categories of personnel employed. Thus, if 80% of the total price is absorbed by junior personnel at, say, a low day rate, then the merits of this can be compared with a situation where 50% of the work is carried out by more qualified/ experienced researchers working at twice this daily rate.

As a general note, to comply with the provisions of Instruction no.1/2017 issued by ANAP, the contracting authority should formulate and apply a quantifiable and easily defendable scoring to each evaluation factor. A good practice is to apply a range of scoring grades based on value-intervals, not on fixed values.

6.2. EVALUATION REPORTS

In some evaluations, especially those that last longer, there is at least one inception report as well as interim report, in addition to a final report. This section provides a brief overview of each.

6.2.1. Inception report

The **Inception Report** which is developed by the contractor must be discussed and agreed with the MA (and the Steering and Scientific Committees if any) in the kick-off meeting following the start of the contract. It will represent, for the whole duration of the exercise, the main point of reference of the quality assurance process (see below), as it states in detail what can be expected from the exercise, the points in time at which the different activities will be performed, and the process through which the evaluation reports will be produced.

The inception report aims to adjust the tender according to the MA's expectations. In particular, the added value of the Inception Report consist of the following: the clarification of the objectives of the evaluation, the finalization of the methodology, the verification of the feasibility of all the evaluation questions, the specification of the operational details (timing, organization of the team, instruments to use for data collection, methods of interactions between the commissioner and the evaluator), clarification of the expected contents of the reports.

In this respect, the inception report sets out:

- the main stakeholders identified;

- the most relevant evaluation questions (elaborated and possibly restated).

- the methods to be employed (methodological approach, analytical techniques, needed data and methods for data collection).

- a detailed work plan with the main tasks and the division of labor between the different members of the team.

- the (finalized) schedule for the work, including the various milestones; and

- the following intermediary and final reports.

6.2.2. Interim and Final Reports

An interim report allows for the sharing of first impressions and provides an opportunity to focus the subsequent stages of an evaluation when early findings highlight such a need.

For draft final reports, it should be emphasized that in the interests of independence, Scientific Committees should concentrate on issues related to the accuracy and conformance to expectations rather than try to second-guess or influence evaluation conclusions.

A possible structure of the final report and the related guiding questions are proposed here:

- **Executive summary** (summarizes the key results and messages; it should be no more than 5-10 pages long and no more than the 10% of the report's length)

- **Introduction** (the context and the objectives of the evaluation, the period when the evaluation was carried out and the main issues tackled in its implementation, a presentation of the structure of the report)

- **Objectives and scope –** (Description of the objective of the evaluation and of the evaluation questions to address, expected use of the evaluation findings; territories and target population covered by the intervention and the evaluation)

- What are the main purposes of the evaluation?

- Who are main users of the evaluation and how they are expected to use it?

- What are the evaluation questions? Who has the evaluation questions defined? Were stakeholders involved in defining the evaluation questions (who and how)?

- What evaluation criteria are involved by the questions?

- What is the scope of the evaluation in terms of interventions, territories and target population covered by the analysis?

- **Background and context** – (description of the evaluated intervention, its strategy in terms of economic and social cohesion, its financial and strategic importance in the OP):

  - Why was the programme/intervention created? What are its main objectives? If only intervention or a group of interventions is evaluated, what is its role in the framework of the overall OP?

  - How was it intended to address the social and economic problems? What is the underlying or explicit ToC?

  - Can other EU or national policies significantly influence the findings of the evaluated programme/intervention? If yes, what are the relevant correlations and complementarities?

  - What are/were the characteristics of its implementation (calls for proposals, eligibility, and selection criteria, allocate resources, etc.)?

  - Does literature review refine the ToC? What findings from other robust evaluations of similar interventions may be used in this evaluation?

  - Were stakeholders involved in the specification of the ToC (who, how, when)?

- **Methodology** – (Methodological approach and its consistency with the evaluation questions and the characteristics of the intervention, fulfilled tasks, used data and data collection methods, data quality and overall reliability)

  - What is the methodological approach and why it is adopted? Is it consistent with the evaluation questions and the characteristics of the programme/intervention (ToC)?

  - How are the evaluation questions addressed? (Types of data used and sources, quantitative

indicators or qualitative descriptors, the analytical techniques, etc.)

- When and how were data collected? What tools were adopted and how scientific standards were respected?

- What are the limitations to the evaluation? (In terms of data availability, methods, and techniques, etc.) How reliable are the findings of the evaluation, especially in the context of the limitations described? To what extent can they be generalized and used for preparing future interventions?

- **Analyses and Results** – (description of the analyses carried out in the study and of the evidence and findings produced by each analysis)

  - What do data and information tell regarding outputs, results, and impacts of the programme/intervention? If relevant, what disaggregation by type of operations, beneficiaries, target groups or territories may be included in the analysis?

  - What are the causes of that results? How are these causes detected?

  - Was the initial ToC confirmed or other explanatory factors come out? Are there unintended or unexpected results?

  - If a mixed method approach is adopted, how do different analyses contribute to provide a consistent picture?

  - What are the limitations of the available findings (e.g., missing data, small sample, no time, or resources for more in-depth analyses, etc.)?

- **Conclusions, lessons learned and good practices** (answers to the evaluation questions and values of the intervention in relation to the evaluation criteria, generalization of the findings, limits of the results for methodological or information constraints lessons learnt and good practices)

  - What are the answers to the evaluation questions? On what evidence are they based on?

  - What is the value of the programme/ intervention in relation to the evaluation criteria covered by the study? How is this value measured?

  - What general conclusions and lessons can be drawn by these findings? Are they clearly and logically rooted in the previous evidence?

  - Is the validity of these conclusions limited by data or methodological shortages? Were

reliability and credibility of the findings discussed with stakeholders and commissioners before completing the report?

- May some good practices be transferred to other programmes or similar interventions?

- **Recommendations –** (policy implications and suggestions for improvements)

  - What are the implications of the conclusions on the evaluated policy?

  - Are there some strategic or implementation deficiencies to recover?

  - What detailed recommendations can be proposed? It is possible to order them by importance and complexity?

  - Who are the actors responsible for implementing the recommendations? What are the resources needed and possible limitations to the implementation of the recommendations?

- **Annexes** (data sets, interviewed persons, documents or statistical tables, instruments used, specifications of the evaluation)

  - Are the annexes in the report easier to read?

  - Are complex and technical methodological aspects included in the annexes and extensively explained here?

  - Are details on questionnaires or interviews shown in the annexes? Are references to the literature included? If useful, are glossary or main definitions included?

A checklist for supporting the control and the assessment of the quality of the evaluation reports is provided in Annex E.

ANNEXURES

7
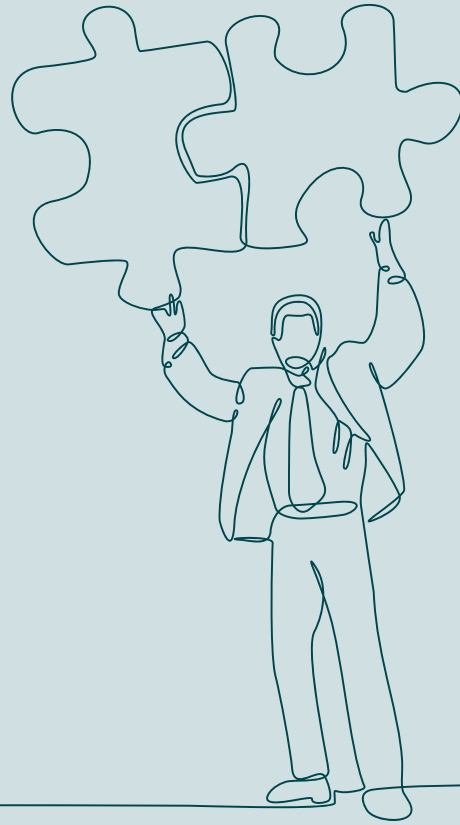
ANNEX A – TEMPLATE OF THE EVALUATION PLAN

This template reflects the indications of the EC's working paper "Performance, monitoring and evaluation of the ERDF, the CF and the JTF in 2021-2027". The template corresponds to an evaluation plan of around 40 pages and uses tables and graphs to summarize contents.

INTRODUCTION

The introduction should include:

- The main objectives of the evaluation plan and its relation with the existing rules (EU Regulations, PA, etc.)

- A short description of the scope of the evaluation plan (which programmes, which Funds and which territories are covered by the evaluation plan and what is the rational on how evaluation plans have been defined in the country)

- A very brief description (around half a page) of the past experiences of the 2014-2020 and the new objectives and challenges in the 2021-2027 period

- The structure and the main contents of the evaluation plan

PART 1 EVALUATION FRAMEWORK

1.1   National governance of the evaluation

This section describes how the evaluation of the ESIF is organized at national level explaining: the role of the different actors (ECU, MAs, evaluation units, evaluation steering committee, evaluation scientific committees (if any, and their national coordination). A graph presenting the organization may help the readers.

This section may be the same, or very similar, in all the evaluation plans. A summary picture like the one presented in the text of the Guide would help present the governance of evaluation.

Guiding questions

How is the evaluation system of ESI funds programmes organized in Romania? How are the competences on evaluation divided between central authorities and MAs of individual OPs? (PA versus OP level evaluations, transversal, or horizontal themes)

- Who are the actors involved and their roles? (ECU, MAs, supporting bodies and committees)

- What are the main coordination functions at national level? A description of the different bodies and committees (or if the case, the selection criteria) should be included, including their management and functioning arrangements

- How is the quality of evaluations ensured throughout the process?

- How is the evaluation capacity promoted at national level? What are the mechanisms and procedures defined to this end?

- How does this evaluation plan interact with the national system (external to ESIF) and how can the role of other key actors be enhanced?
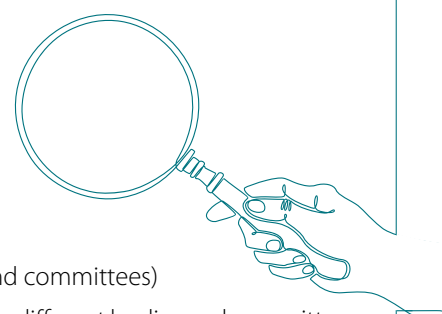
1.2 Coordination and ORGANIZATION of the evaluation at the OP level

This section introduces the organization of the evaluation at the OP level: role and responsibility of the different actors (MA, evaluation unit, Monitoring Committee, Evaluation Steering Committee, Evaluation Working Group set up at national level, implementers, experts to support the administration, independent evaluators). A graph presenting the organization may help the readers.

Guiding questions

- How is the evaluation of this OP (or these OPs) organized?
- Who are the main actors and their roles? What are the main competences on evaluation of the Monitoring Committee and the different supporting groups and committees (Evaluation Steering Committee, Scientific Committee etc.)?
- How is the coordination between the different actors guaranteed? How do the responsible offices for evaluation interact with the implementers of the OP and the other actors (including those from other OPs and from national level, including in the context of the different groups and networks organized)?
- How is the quality of evaluations assured?
- How is the evaluation capacity promoted at the OP (or OPs) level?
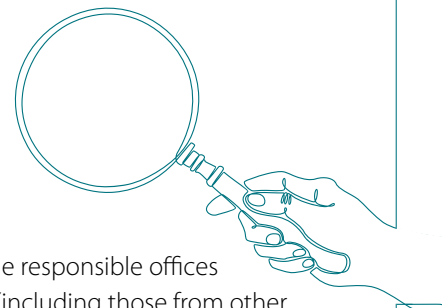
1.3 Involvement of the stakeholders in the evaluation plan

Who and how have the stakeholders been involved during the preparation of the plan and will be involved during the plan's implementation?

The role of the stakeholders in the management and implementation of the evaluations (consultation in the definition of the evaluation questions and in the use of the results, active participation in some evaluations because directly involved in the interventions, information and dissemination of results, training on evaluation) and how the participation of the stakeholders is ensured and promoted (monitoring committee, evaluation group).

A table summarizing the main interactions with the stakeholders in the preparation of the plan should be included in the report. An example of the table is below.

Method of involvement[1] and date	Involved stakeholders	Main results of the discussion and main collected information

(1) meetings, seminars, surveys, etc.

Guiding questions

- What are the stakeholders of the OP (OPs)? How are they divided by the level of influence on the OP?
- How were they involved in the preparation of the Evaluation Plan? What contribution did they provide?
- How is it planned to involve the stakeholders during the implementation of the evaluation plan? Are the involvement of some categories of stakeholders more frequent and why?
- How will these stakeholders be involved in the use of results? Are they direct users or do they play a role in an enhanced use of results?
- What are the venues and the methods to discuss with stakeholders? On what aspects of the future evaluations their contribution is considered fundamental?
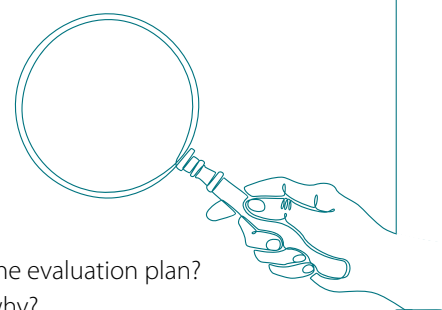
1.4   Strategic approach to the evaluation

This section explains how the evaluations of the plan have been identified and what are the most important and strategic elements of the plan. The following four points can be subsections of this section and for each of them some guiding questions and tables (where useful) are suggested:

1.4.1   The strategy of the OP and the main components to evaluate based on the ToC or IL
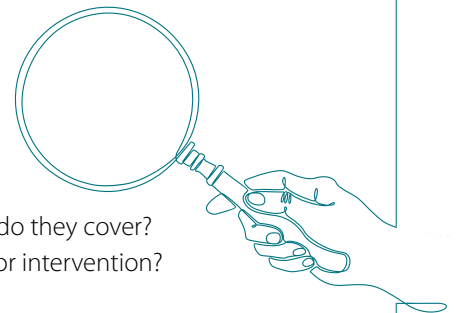
Guiding questions

- What are the main objectives of the plan? On what components of the OP does the evaluation plan focus on?
- How is this evaluation strategy justified in comparison to the OP (or OPs) strategy? Is this approach confirmed by the ToC of the OP (OPs)?
- How are the implementation and the impact issues balanced in the overall evaluation strategy? What does the past experience suggest on this respect?

1.4.2   Key evaluation questions deriving from national and EU strategies

Guiding questions

- What are the key evaluation questions deriving from the EU and the national strategies relevant for the OP?
- How were these questions defined?
- How these questions are related to the ToC and what components of the OP do they cover? How do they relate to the more direct questions concerning the OP strategy or intervention?
- What are the evaluation criteria involved by these evaluation question?
- What are the other minor questions that can be clustered around each of the key questions?

1.4.3   Collection of other evaluation questions by implementers and stakeholders

Guiding questions

- What are the other key evaluation questions?
- Do these questions derive from the discussion with stakeholders or the internal reflections of the MA?
- How do these questions relate to the ToC and what components of the OP do they cover?
- What are the evaluation criteria involved by these evaluation questions?
- What are the other minor questions that can be clustered around each of those key question?
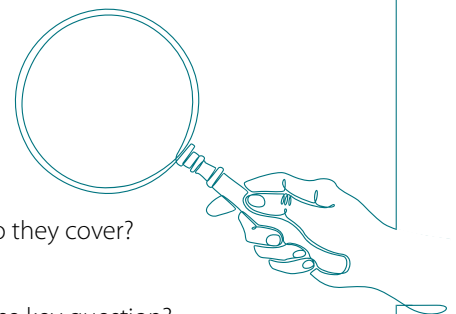
To summarize the evaluation questions, the following table, or a similar one, can be included in a unique frame of all the main relevant items.

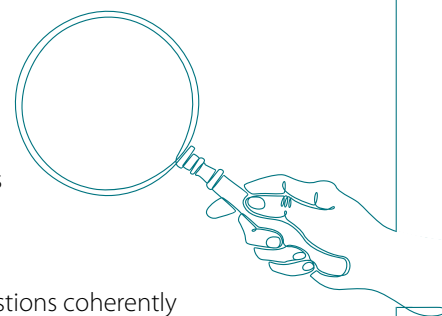| Key evaluation question	How key question was identified[1]	What component(s) of the OP is covered by the question[2]	What evaluation criteria[3] are related to the question	Secondary questions related to the key question[4]

Notes:

1) *in general main methods are: a) descending from EU or national strategic priorities, b) demanded by the stakeholders or the MC, c) requested by MA or implementers.*

2) *The question can refer to a specific measure, to a part of the OP or to the entire OP. Specifying the scope helps to understand needed data and methodology.*

3) *Effectiveness, efficiency, relevance, coherence, EU added value, sustainability or others key criteria, A question generally refers to only one criterion.*

4) *As mentioned above, these questions generally are specifications of the key question.*

1.4.4 Main contents of the evaluations and their methodological approaches (summary of the types of planned evaluations, main approaches and objectives of the evaluations)

Guiding questions

When this sub-section is drafted, the part 2 of the evaluation plan, which details the planned evaluations (see below), must be completed. In brief, this section summarizes part 2 and does not add any original elements, but it only introduces the planned evaluations

- What evaluations have been planned?

- What are the key questions associated with each evaluation? Are the key questions coherently assembled for each evaluation and is that the right number of questions (i.e. not too many questions)?

- What are the evaluations focused on the implementation of the OP and what are those focused on its results and impacts?

- Are the necessary data and timing of the planned evaluation sufficiently defined? Are there any feasibility problems for any of the evaluations? And, if so, do these evaluations need some more in-depth examination in the future?

- Do the planned evaluations need some preparatory studies or agreements with other institutions to provide necessary data?

In the subsection, a table (see below) could support the presentation of the planned evaluations and their main characteristics.

N. or code	Planned evaluation (title)	Associated key evaluation questions	Period (End date)	Type of evaluation[1]

Note: 1) possible type of evaluation or studies are: a) Preparatory study, b) data provision, c) Implementation or process evaluation, d) Impact evaluation

1.5   Evaluation implementation and quality

This section describes how the evaluations will be implemented and the main factors that will ensure their quality:

• Role played by each actor in the process (MA, Evaluation Scientific Committee, Evaluation Steering Committee, Monitoring Committee)

• Independence of the evaluator

• External evaluations and call for tender to identify the external evaluators

• Call for tender quality guaranteed by: structure of the terms of references and main criteria to select the

evaluators (independence, technical capacity, quality of the methodological proposal)

• Data availability (main data to be used in the evaluations; commitment of the MA to make available administrative data, if necessary; data issues to tackle in the evaluations)

• Quality insurance of the process: quality checks during the implementation (intermediate reports and meetings); check of the final report with the support of experts, workshop to discuss and validate methodology, used data and results.

Guiding questions

• Has a strategy to assure evaluation quality and reduce potential risks in the evaluation process been activated? What are its main elements?

• How will the different actors contribute to ensure evaluation quality?

• How is the independence of the evaluators guaranteed?

• How are the call for tenders for evaluators organized to ensure high-quality evaluations (including aspects on selection criteria)? How is the quality of the proposals measured? What solutions are put in action to respect the planned timing of the calls for tenders?

• How is the implementation of the evaluations supervised to guarantee quality? How are the necessary skills to check evaluation quality ensured?

• How is data availability checked, promoted and ensured when directly depending on MA? Are agreements with data owners or other initiatives to easily accessing data promoted/ signed?

• How is the quality of the evaluation reports checked? (instruments, procedures, resources, timing)

• Do any of the steps of the evaluation process need particular attention to ensure evaluation quality? If yes, why and what are the measures planned to this end?
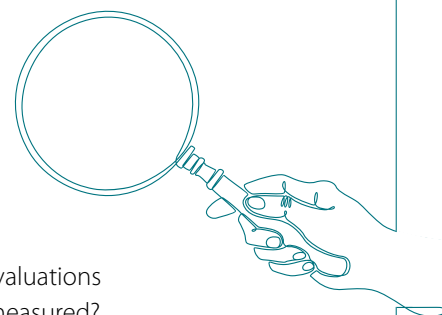
1.6   Dissemination and use of the findings

This section introduces the strategy to diffuse results and use findings of the evaluations:

- Objectives of the dissemination of the results and targets to reach out (political and administrative policy-makers, stakeholders and beneficiaries, researchers)

- Role of the Monitoring Committee and other venues to discuss results of the evaluations (evaluation groups, workshops with stakeholders, specific presentations to the policy makers and political decision-makers)

Guiding questions

- Who are the main users of the evaluation findings? How will they be involved in the preparation of the evaluations and in the discussion of their findings?

- Are users and stakeholders classifiable in different categories and involved in different ways? Are specific tools or places defined to involve different categories of users?

- Is a standard process of dissemination of findings planned for all the evaluations? What are the main dissemination methods and how will these be adapted to the characteristics of each category of (potential) users of evaluation results? Are there some pre-defined venues (e.g., MC or Evaluation Group) devoted to the discussion of the evaluation findings?

- What is the mechanism foreseen for monitoring the implementation of evaluation recommendations? How is this organized and what resources are needed?
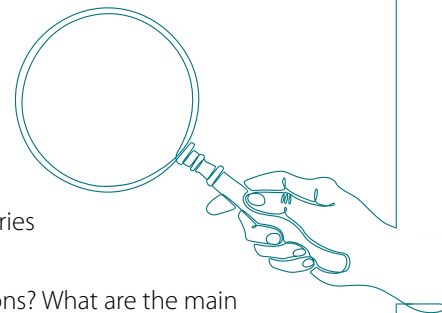
1.7   Evaluation capacity building

This section describes the planned initiatives to reinforce evaluation capacity in the context where the OP is implemented. Capacity building may be directed to:
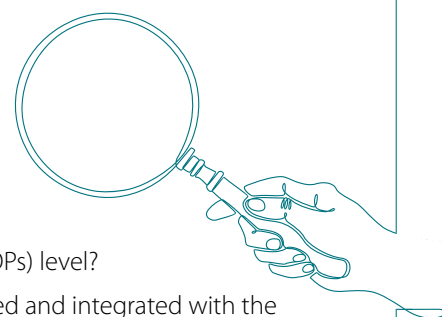
- The MA and the administration (e.g., training and technical support to the evaluation unit, the other offices of the MA and the offices responsible for the implementation)

- Committees and other bodies set-up to support the evaluation process and promote a coordinated approach of evaluation, or facilitate the dissemination

of evaluation results (e.g., Evaluation Group, Evaluation Coordination Committees, members of Monitoring Committees or CCMAP etc).

- The external environment (e.g., training and workshops with the stakeholders, market analysis and exchanges through meetings with the evaluation companies, training or seminars with the municipalities or other institutions to prepare and discuss evaluation, workshops with academic experts or Evaluation society).

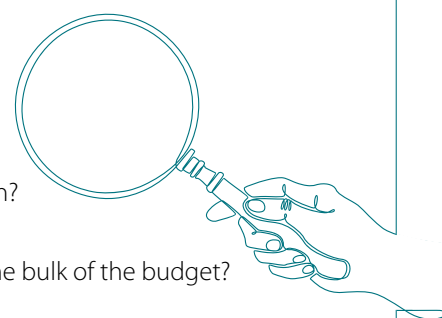## 1.8   Available resources and timing of the evaluation

This section summarizes the estimated financial resources for implementing the evaluation plan and the period of their expenditure. The following table provides an example of how this may be presented including all the activities (evaluations, technical support, studies, capacity building, etc.) of the evaluation plan.

Evaluation functions and main activities	Duration (period from … to ….)	Estimated cost (Euro)	Financial sources
Technical support and coordination of the MA and the evaluation unit			*(Priority and Specific Objective of the OP)*
Preparatory studies and data provision			
Evaluation studies			
Dissemination of results and events			
Capacity building initiatives			

PART 2 PLAN OF THE EVALUATIONS

This second part of the plan provides details on the individual evaluations and consists of two sections:

1. First section provides a synthetic overview of all the planned evaluations through a brief introduction and a summary table.

2. Second section contains the illustrative fiches of all the planned evaluations.

2.1   Lists and timetable of the evaluations

Table to summarize planned evaluations

N. or code	Objective of the evaluation	Content and scope of the evaluation			Period	Type of evaluation	Planned Cost
		Priority Axis(es)	Specific Objective (s)	Interventions			
	(Max 50 words)			(max.100 words)	Initial date -end date)	E.g., - Preparatory study, - data provision, - Implementation (or process) evaluation, - Impact evaluation	(Euro)

Guiding question

• Are all the sections of the table completed and clearly understandable to all readers?
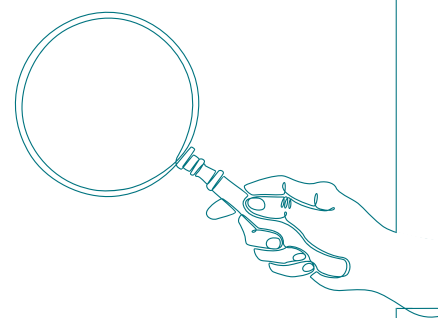
2.2   Fiches of the planned evaluations

Template of the fiche of a planned evaluation

Evaluation n. or code – Title……………	
Priority and specific objectives covered by the evaluation	(Priority axis, specific objectives)
Types of interventions to be evaluated	(Short description of the policy measures and state aids evaluated, when identifiable) 100 words
Type of evaluation	(e.g., preparatory study, implementation evaluation, impact evaluations)
Focus and rationale of the evaluation	(Main focus and objectives of the evaluation, its expected use, main characteristics of the intervention to evaluate, etc.) 300-500 words
When the evaluation will be implemented	(Period from …. To …….)
Main evaluation questions	(List of the key evaluation questions, whit some short description of the expected analysis for each of them) 200-250 words

Evaluation n. or code – Title..............	
Methodological approach and possible methods	(Statistical, modelling, counterfactual, theory based, qualitative research, etc. and indications of the main methods to be probably used)
Data sources	(Used data among primary secondary and administrative data; main types of data to use, possible arrangements to facilitate collection and access to data)
How the evaluation will be implemented	(Internal or external evaluator, what type of call for proposal: open / restricted, framework contract, one or more lots, etc.)
Quality insurance	(How the MA will ensure the quality of the evaluation)
Planned cost (Euro)	(Planned cost)

Guiding questions

- Are all the sections of the fiche completed and clearly understandable to unexpert readers?
- Are sufficient details of the most complex aspects (e.g., methodology) provided?
- Are the different aspects of the fiche well consistent in each evaluation? (e.g., methodology with the type and complexity of evaluation and the evaluation questions)
- Are the fiches coherent with the previous paragraphs of the plan ?
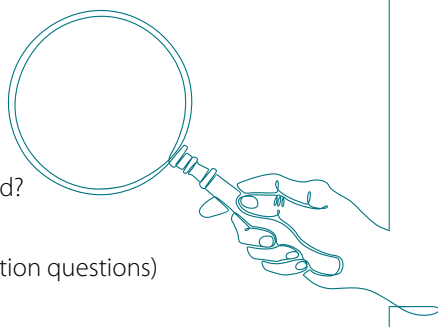
ANNEX B – TOOLS TO ORDER AND CLASSIFY EVALUATION QUESTIONS

CHECKLIST FOR ASSESSING YOUR EVALUATION QUESTIONS

Created by CDC's National Asthma Control Program 2013

The success of an evaluation depends on creating appropriate overarching evaluation questions. Once you have drafted a set of potential evaluation questions, apply the criteria below to each question. Reviewing the questions may help you to identify the ones that are most likely to provide useful information. Although no set criteria can be universally applicable, this checklist should be helpful regardless of the purpose of your evaluation.

Does the evaluation question meet this criterion?	YES	NO	Does not meet criterion but merits inclusion because...
QI.			
1. Stakeholder engagement			
A. Diverse stakeholders, including those who can act on evaluation findings and those who will be affected by such actions (e.g., clients, staff), were engaged in developing the question.			
B. The stakeholders are committed to answering the question through an evaluation process and using the results.			
2. Appropriate fit			
A. The question is congruent with the program's theory of change.			
B. The question can be explicitly linked to program goals and objectives.			
C. The program's values are reflected in the question			
D. The question is appropriate for the program's stage of development.			
3. Relevance			
A. The question clearly reflects the stated purpose of the evaluation.			
B. Answering the question will provide information that will be useful to at least one stakeholder.			
C. Evaluation is the best way to answer this question, rather than some other (non-evaluative) process.			
4. Feasibility			
A. It is possible to obtain an answer to the question ethically and respectfully.		Unless an acceptable option can be found, eliminate this question.	
B. Information to answer the question can be obtained with a level of accuracy acceptable to the stakeholders.			
C. Sufficient resources, including staff, money, expertise, and time can be allocated to answer the question.			
D. The question will provide enough information to be worth the effort required to answer it.			
E. The question can be answered in a timely manner, i.e., before any decisions potentially influenced by the information will be made.			

Does the evaluation question meet this criterion?	YES	NO	Does not meet criterion but merits inclusion because...
5. In sum ...			
A. This question, in combination with the other questions proposed for this evaluation, provides a complete (enough) picture of the program.			
B. The question, in combination with the other questions proposed for this evaluation, provides enough information for stakeholders to act.			

WORKSHEET 5:

PRIORITIZE AND ELIMINATE QUESTIONS

Take each question and apply the criteria below:

Question	Which stakeholders?	Importance to stakeholders	New data collection?	Resources required?	Timeframe	Priority (High, Medium, Low or Eliminate)
						H M L E
						H M L E
						H M L E

ANNEX C – PROMOTING A CLEAR AND EASY WAY TO COMPLY WITH DATA PROTECTION RULES [34]

Difficulties can be experienced in obtaining data that identify individuals or companies who have participated to financed interventions. For instance, Counterfactual Impact Evaluations (CIE) require micro-data - data which contain observations on the individual units in both treatment and control groups.

ESF+ Regulation 2021/1057 (Annex I)) asks for data on participants with a breakdown by gender, labor market status, age group, educational attainment, and vulnerable groups (migrants, minorities, disabled, other disadvantaged). The CPR and ESF+ Regulations for 2021 - 2027 establish a legal obligation for MAs to collect and process personal data in the form of individual participant records. In addition, ESF+ Regulation 2021/1057 at Article 17(6) in relation to Monitoring and Indicators specifies that: *"Where data are available in registers or equivalent sources, Member States may enable the managing authorities and other bodies entrusted with data collection necessary for the monitoring and the evaluation of general support from the ESF+ strand under shared management to obtain data from those registers or equivalent sources, in accordance with Article 6(1), points (c) and (e), of (EU) 2016/679".*

These rules, set out in the Common Provisions and ESF+ Regulations, facilitate access to and use of personal data needed for ESF+ monitoring and, in the case of a CIE, for defining the treated group. However, access to personal data necessary to evaluation must be in line with Regulation (EU) 2016/679 (General Data Protection Regulation – GDPR), which covers the general transfer and use of personal data, including special categories of data[35] within the EU.

Relevant legislation on the processing of personal data in Europe consists mainly of Regulation (EU) 2016/679 and the guidelines and measures adopted by competent authorities such as the European Data Protection Board (EDPB). This legislation sets out many conditions and limitations on the processing of personal data to protect the rights and freedoms of data subjects. At all events, the need for protection and safeguard that emerges from the provisions and regulations must find a balance with the need not to constrain scientific research and indeed to act as an asset in its development. For this reason, and within this legal framework, the European Union has foreseen that, under specific conditions, exceptions can be made to allow research activities and the dissemination of the outputs if the first and foremost right of the persons concerned, i.e., the right to privacy, is safeguarded.

GENERAL DATA PROTECTION REGULATION (EU) 2016/679

The General Data Protection Regulation (GDPR) came into effect on May 24, 2016 and became fully applicable in all Member States on May 25, 2018.The GDPR applies "to the processing of personal *data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system".*

For processing personal data, the following main principals set out in Article 5 of the Regulation is essential to know:

a. *Lawfulness, fairness, and transparency:* personal data shall be processed lawfully, fairly and in a transparent manner.

b. *Purpose limitation:* personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes. In the case of processing for statistical purposes or scientific research, the data controller shall adopt the necessary guarantee and protection measures.

c. *Data minimization:* only data strictly necessary to achieve specific purposes should be collected.

d. *Accuracy:* the data collected shall be accurate and, where necessary, kept up to date.

e. *Storage limitation:* personal data shall be kept in a form that permits identification of data subjects

for the appropriate amount of time. To this aim, personal data may be stored for longer periods as the data is processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89. In such cases, the data shall be subject to the implementation of appropriate technical and organizational measures to safeguard the rights and freedoms of the data subjects.

f. *Integrity and confidentiality:* personal data shall be processed in a manner that ensures appropriate security thereof, including protection against

unauthorized or unlawful processing and accidental loss, destruction, or damage, using appropriate technical or organizational measures.

g. *Accountability:* one of the most important principles set out by the Regulation. It states that the controller is responsible for data processing and can demonstrate the implementation of any required measures.

This set of principles constitutes the main structure that GDPR is based on and determines a series of obligations to be fulfilled by the controller and the processor.

PROCESSING OF PERSONAL DATA IN THE CONTEXT OF THE ESIF EVALUATIONS: CONDITIONS, LIMITS AND MAIN ISSUES

Conducting an evaluation often involves processing large amounts of data, including personal data. Referring to the regulatory framework described above, it is important to be familiar with the conditions, limitations, and main issues that MAs face when conducting an evaluation.

Legal basis and purposes of the processing

To carry out the processing of personal data, at least one of the conditions of lawfulness indicated in Article 6 of the GDPR must be present[36]. If, on the other hand, the data to be processed are "special categories of data", a reference must also be made to Article 9 GDPR.

In the context of the ESIF evaluation, **public interest** according to art.6(1) e) of the GDPR (*"the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller"*) appears the most appropriate legal basis for data processing. This legal basis must be laid down in EU or national law, as specified in art. 6(3) of the GDPR. The 'public interest' clearly represents the obligations of the managing authorities defined in ESF+ Regulation 2021/1057 art.17(6) on the use of data *"available in registers and equivalent sources"*. In addition, CPR 2021/1060 states in art. 4 that *"the Member States and the Commission shall be allowed to process personal data only where necessary for the purpose of carrying out their respective obligations under this Regulation, in particular for monitoring, reporting, communication, publication, evaluation [...]"*. National laws can also vest data controllers with similar authorities for managing and processing data in the public interest.

Other legal bases for an evaluation may be referred to in other conditions listed in art.6(1) and in particular:

•   the **consent** of the data subject (art.6 (1) a) GDPR). The consent, for instance, may be an appropriate legal basis when data for the CIE are collected by a survey and the data subjects can easily give their consent to the processing (see art.7 of GDPR on consent). In general, consent is more complex if not planned well in advance; it may be considered a "residual" legal basis for evaluation when other legal bases are not applicable.

•   processing is necessary for **compliance with a legal obligation** to which the controller is subject (art.6 GDPR in reference to art.17 of ESF+ Regulation). This legal basis must be established by law and may involve private or public entities; it could be, for instance, due to specific legal prescriptions, a private or public entity responsible for a dataset is required to collaborate with the managing authority in the evaluation.

In addition, the data controller **may use the collected data for further purposes if these are compatible with the initial purposes**. In this regard, the data controller will have to evaluate the conditions laid down in Articles 6(4) and 5(1) b) of the GDPR. These provisions are particularly relevant when processing data for scientific research or statistical purposes. They allow the use of administrative data for purposes that differ from the original ones and do not require specific consent when repurposing the data

---

36 These conditions are: a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes; (b) processing is necessary for the fulfillment of a contract to which the data subject is a part of or in order to take steps at the request of the data subject prior to entering into such contract; (c) processing is necessary for compliance with a legal obligation to which the controller is subject; (d) processing is necessary in order to protect the vital interests of the data subject or another individual; (e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller; (f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the   data subject which require protection of personal data, in particular where the data subject is a minor.

but are still required to comply with the protection rules specified in art. 89(1), mainly pseudo-anonymization (see below). These provisions, for instance, may be relevant in the case of unemployment register data to be used in a CIE.

Due to the specificity of evaluations, the data controller collecting the data and the entity which carries out the research may be different. In this case, there must be a condition that legitimizes the data transfer and allows the recipient to proceed with the evaluation.

Anonymization and pseudonymization: processing for statistical purposes

Article 89 of the GDPR states that data processing carried out for purposes of public interest, or in the context of scientific research or for statistical purposes, shall provide for appropriate safeguards for the rights and freedoms of data subjects and shall respect the principle of "minimization". This means making use of pseudonymization techniques[37].

Where the purposes can be fulfilled by subsequent processing operations which do not permit, or no longer permit, the identification of data subjects, those

purposes shall be fulfilled by anonymizing the data and then processing them in aggregated form. In many evaluations, results are aggregated, and this risk is absent, unless the original datasets are published for scientific reasons. In this case, datasets must be anonymized.

Data storage and secure processing

One of the main aspects of data processing is data storage. The legislation does not specify how data must be stored, but the principles mentioned above require that storage and processing should always be linked to the purpose of the research. When the purpose of the processing is achieved, retaining the data is no longer necessary and can be disposed. This general rule must be specified in the privacy statement that is given to data subjects when they are registered in administrative datasets or, given the likely impossibility of informing all data subjects personally, alternative ways of providing the information may be found (for instance, by publishing an information page containing the privacy policy on the research activities on the website of the managing authority).

LEGAL OBLIGATIONS

Information for data subjects

The first duty of the data controller[38] is to inform the data subjects. This obligation is set out in Article 13 GDPR when the data is collected from the data subject; other cases are regulated by the provisions under Article 14. The subject must be informed about the purposes and methods of the processing, the legal basis, the storage periods, and the rights that can be exercised by the data subject. In the case of ESIF, this information is given at the time of collecting the data which will be later used for statistical and evaluation research, such as in ESF or ERDF monitoring or unemployment register datasets. When this communication is not possible, especially in the case of individuals in registers used for the evaluation, a website or other general tools of information can be used.

Governance of the relationship between the different entities involved in processing

Generally, research activities involve more than one body and, in such cases, the relationships between them must be regulated by specific arrangements known as "data

processing agreements". These are to be defined on a case-by-case basis reflecting the contributions of the different stakeholders. Examples of possible relationships are:

- *Data controller – Data controller* occurs when the entities collaborate in the implementation of a project, although in different conditions and with different tasks, each maintaining its own distinct processing purpose.

- *Data controller - Data processor* occurs when one entity (Data controller) determines the means and purposes of processing and uses another entity (data processor) to perform certain processing activities. In this case, the reference and obligations are to be found in Article 28 GDPR.

- *Joint controllers:* this type of relationship is set out in Article 26 of GDPR, which provides that: *"Where two or more controllers jointly determine the purposes and means of processing, they shall be joint controllers".* In such cases, the parties must determine their

---

37 Pseudonymization means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific person without the use of additional information. Anonymization refers to the processing of personal data in a manner that makes it impossible to identify individuals from them. An overview of pseudonymization techniques can be found at https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices.

38 The entity which is collecting the information. In general, it is the MA.

respective responsibilities for compliance with the regulations in a transparent manner, in particular regarding the rights of the data subjects.

Data protection impact assessment

Article 35 of the GDPR states that, where the processing of personal data is likely to present a potentially high risk to the rights and freedoms of data subjects, the controller must carry out an impact assessment before proceeding with the data processing. An impact assessment in evaluations may be important only in case of processing special categories of data subjects on a large scale. In such cases, the data controller in the administration may conduct the impact assessment according to the rules and the tools made available by the national Data Protection Authority.

FUTURE PROSPECTS: THE DATA GOVERNANCE ACT

The framework outlined above reflects the current state of data protection and the rules that are in place in the EU. The GDPR was introduced specifically to enable a breakthrough in personal data protection, and to create a common system across the EU in accordance with current technological developments and today's use of data in society. Other data reforms are expected to be made and may have significant impact on the collection of personal data in the coming years.

The Data Governance Act is currently still at the proposal stage in the European Commission[39]. The explicit aim of this act is to promote the availability of usable data by strengthening trust in data intermediaries and enhancing data sharing mechanisms across the EU. Personal data are likely to be the subject of the new standard, and their use for statistical and research purposes could be far more reaching thanks to the introduction of a new figure: the personal data-sharing intermediary. Such a measure could contribute to creating a more enabling environment for CIEs, allowing easier access to a large amount of information on a large scale.

---

39   The proposal for a Data Governance Act was adopted at the end of 2020 by the Commission. The state of play of the Data Governance Act can be found here: https://eur-lex.europa.eu/legal-content/EN/HIS/?uri=CELEX:52020PC0767.

ANNEX D – CHECKLIST OF THE TERMS OF REFERENCES

This checklist supports the definition of the Terms of Reference (ToR) for identifying an evaluator; it is designed to verify the presence in the ToR of all the needed elements and the pertinence of the ToR.

The checklist adopts a general approach and for this reason is applicable to different types of evaluations and ToRs. For the same reason, it may happen that in specific cases, some additional items may be necessary (e.g., in counterfactual impact evaluation, a reference to the identification of the control group may be useful); however, the additional items can be easily derived from the proposed items (for instance, concerning the control group see questions 3.2.2. and 3.2.3.). According to habits and prevailing rules, some items may be placed in the administrative specifications or in the technical specifications (e.g., price and payments); this different location does not affect the use of the checklist.

A positive answer to the questions of the checklist indicates that item is not problematic, the opposite for the negative answers.

Checklist	Y/N
The selection procedure	
1.1. Is the procedure to select the evaluator compatible with the timing of the planned evaluation?	
1.2. Does this procedure provide a good balance between the large access of evaluators (competitiveness) and the time needed for the selection (efficiency)? Is this procedure more effective/efficient than other procedures of public procurement?	
1.3. Has the office charged for managing the procedure the skills and the personnel sufficient to successfully implementing it?	
1. The administrative specifications	
1.1. Are all the key elements of the procedure included in the administrative specifications (e.g., functioning of the procedures, deadlines, criteria to participate and be selected, etc.)?	
1.2. Are the main deadlines specified (e.g., to request additional documentation, to pose questions, to submit the tender)?	
1.3. Are the eligibility criteria to have access to the call for proposals clear, in line with the national and EU rules and do not create serious limitations to competition?	
1.4. Are the selection (or quality) criteria clear and capable to identify the best quality proposal?	
1.5. Is the weight of price in comparison to the other selection criteria balanced and not excessive?	
1.6. Is the way to apply (interpretation, scoring, ranking, etc.) the selection and price criteria specified?	
1.7. Are the composition and role of the selection committee defined?	
1.8. Is the documentation to submit clearly identified and does it include standardized application forms or other tools to simplify and minimize errors?	
1.9. Is the structure of the technical offer indicated (main contents, chapters, length, etc.)?	
2. The technical specifications	
2.1. Context, objectives and scope	
2.1.1. Is the policy context of the evaluation (EU regulation, evaluation plan, OP and other EU or national relevant decisions) explained?	
2.1.2. Are the main objectives and the users of the evaluation identified?	
2.1.3. What type of evaluation (e.g., preliminary study, implementation or process, impact, mix of different types) is defined?	

Checklist	Y/N
2.1.4. Are the interventions to evaluate, the territory to cover and the period to examine (the scope of the evaluation) well-defined and clearly distinguishable?	
2.1.5. Is a brief description of the implementation and the advancement of the interventions to evaluate provided?	
2.1.6. Are the key stakeholders of the evaluation identified?	
2.1.7. Are the evaluation questions clearly stated? Are the key evaluation questions well-defined?	
2.1.8. Is the ToC of the interventions to evaluate clarified? Or is the evaluator requested to identify the pertinent ToC?	
2.2. Methodology	
2.2.1. Is the general methodological framework suggested? And is there a request for major specification of the methodological approach made?	
2.2.2. Is the data expected to use clearly defined? And is a request made for major specification of necessary data and collection tools?	
2.2.3. Is a request for clarifying the main methodological techniques and analyses to use clearly made?	
2.2.4. Are the main tasks to fulfill in the evaluation identified?	
2.2.5. Is there a request for specifying the methods used to validate results and findings of the evaluation made?	
2.2.6. Are the main deliverables (reports, meetings) of the evaluation defined? And are their main expected contents specified?	
2.2.7. Are a risk assessment of the evaluation process and a specific quality control requested?	
2.3. Professional qualifications	
2.3.1. Are requirements for skills and experience of the team clearly defined? And are these requirements coherent with the service requested?	
2.3.2. Are requirements for skills and experience clearly interpretable, sufficiently wide to be found in the market and not limit competition?	
2.3.3. Is the multidisciplinary composition of the team expressly detailed (if necessary)?	
2.3.4. Is the request of specifying the distributions of roles and responsibilities in the team made?	
3. Budget and Payment	
3.1. Is the maximum price for the evaluation stated?	
3.2. Is specified how the budget of the evaluation must be presented (total cost, detailed budget for main voices, etc.)?	
3.3. Are the timing and the amount of the payments unambiguously defined?	
4. General	
4.1. Is the number of objectives and evaluation questions not excessive? Can they be addressed in a unique evaluation?	
4.2. If doubts on the feasibility of the evaluation exist, is a feasibility analysis included in the requests and a potential "plan B" defined (e.g., alternative approaches or the break of the contract)?	
4.3. Is the language used clear, simple, and always well-focused on the main elements?	
4.4. Are all the requests sufficient and adequate to assess the proposals according to the adopted selection criteria?	

ANNEX E – CHECKLIST FOR ASSESSING THE EVALUATION REPORT

The Evaluation Report Checklist[40] has two intended purposes that are related to evaluation management: (1) a tool to assess the content of the evaluation report by the Evaluation Commissioners and (2) a tool to guide evaluators, when preparing the report. Evaluators can self-rate their own progress during the writing phase. They can also use the checklist to identify weaknesses or areas that need to be addressed in their evaluation report(s).

Instructions: Rate each component of the report using the following rubrics. Place a check mark in the cell that corresponds to your rating on each checkpoint. If the item or checkpoint is not applicable to the report, indicate the "NA" cell to the far right.

1=Not addressed, 2=Partially addressed, 3=Fully addressed, NA=Not applicable

	1	2	3	NA
1. Executive Summary				
1.1. The programme/ IP/ SO/ theme evaluated is well described				
1.2. Evaluation questions and purpose of the evaluation are presented				
1.3. A brief description of methods and analytical strategy (if appropriate) is provided				
1.4. A summary of main findings and policy implications or recommendations is included				
1.5. Length is adequate (in general no more than 10-12 pages, or around 10% of the report)				
1.6. Comments:				
2. Introduction				
2.1. The introduction helps the reader in approaching the report				
2.2. An overview of the report and the description of report structure are available				
2.3. Objectives and scope of the evaluation are clearly presented				
2.4. The programme/ intervention to evaluate, its expected use and relevant users are specified				
2.5. References of the evaluation to the Evaluation Plan and other possible decisions of the MC are included				
2.6. Evaluation questions and how they have been identified (e.g., interviews, surveys, discussion with the MA, meetings with MC and the stakeholders, etc.) are clearly described				
2.7. Evaluation criteria included in the analysis are specified, as well as their relations with the evaluation questions				
2.8. The target population of the programme/ IP/ SO (as relevant) and territorial areas covered by the intervention are clearly identified				
2.9. The main stakeholders of the evaluation are clearly identified				
2.10. Comments:				
3. Background and context				
3.1. A description of the programme/ IP/ SO/ theme being evaluated (its strategy in terms of economic and social cohesion, strategic importance in the OP, etc.) is included				
3.2. The cause-effect relations underlying the programme/intervention are explicitly presented (a ToC or other interpretative framework)				
3.3. The implementation of the programme/ intervention is well described and allows to understand possible bottlenecks or difficulties				

---

40 The checklist uses different sources and adapts their contents according to the experience of the authors; see: Evaluation Checklist, Gary Miron (2004); Checklist for preparing the Evaluation Report ILO (2021); EVALSED: The resource for the evaluation of Socio-Economic Development (2013)

	1	2	3	NA
3.4. The main interactions with other relevant European or national policies are identified and described				
3.5. A well-focused review of the related literature is available to identify what is already known (including. aspects on previous and similar financing and lessons learned etc.)				
3.6. Comments:				
4. Methodology				
4.1. Evaluation approach and its rationale are clearly described and fit the ToC and the evaluation questions				
4.2. Sources of information and data are adequately presented (e.g., primary, or secondary data, sampling method, statistical error, questionnaires, timing of data collection, etc.)				
4.3. Analytical techniques are well described and allow to understand the reliability of the results				
4.4. The strategy of combining methods/approaches (if any) is justified and allows to answer the evaluation questions properly.				
4.5. Possible limitations of the evaluation are specified (e.g., limitations related to methods, data sources, potential sources of bias etc.)				
4.6. Comments:				
5. Main findings				
5.1. The methodology is correctly applied				
5.2. Details of analyses and findings are clearly and logically described.				
5.3. Analyses and findings cover all main aspects as deriving from the cause-effect relationships identified with the help of the ToC or other interpretative framework used				
5.4. Discussion of evaluation findings is objective and complete, including – where relevant – both negative and positive findings				
5.5. Findings are supported by evidence and are consistent with methods and data used				
5.6. All evaluation questions are addressed, and an explanation is included for questions that could not be answered				
5.7. Findings with regard to the examined evaluation criteria and the evaluation questions are presented				
5.8. Unintended and unexpected results are discussed (if the case, applying to impact evaluations)				
5.9. Factors contributing to the success/failure of the programme /intervention are identified and discussed				
5.10. Comments:				
6. Conclusions, lessons learned and emerging good practices				
6.1. Answers to all evaluation questions and values of interventions/ themes in relation to the evaluation criteria are provided				
6.2. Conclusions are formulated by synthesizing the main findings into summary judgments of merit and worth (any limitations of the results should be also explained)				
6.3. Conclusions are fair, impartial, and consistent with the findings				
6.4. Conclusions are clear, concise and their potential generalization (at the level of a larger target groups, in time or in the space) is clarified				
6.5. Conclusions reflect the analysis of horizontal or cross-cutting themes (including transterritorial relationships in ETC, gender and environmental sustainability) conducted in the evaluation				
6.6. Lessons learned, including context and applicability are included (if the case)				
6.7. Emerging best practices, including context and applicability are included (if the case)				
6.8. Comments:				

	1	2	3	NA
7. Recommendations and policy implications				
7.1. Recommendations logically follow from findings and conclusions, lessons learned and good practices				
7.2. Recommendations indicate the action needed to improve the performance of the programme/intervention or the institutional set-up in a concise manner. Long sentences and paragraphs are avoided				
7.3. Recommendations are sufficiently detailed (who is called upon to act, time frame for their implementation, costs and/or complexity, etc.)				
7.4. Recommendations were discussed and validated with implementers and stakeholders (if requested or useful)				
7.5. Comments:				
8. Annexes and references				
8.1. A suitable style or format is used consistently for all references				
8.2. Annexes included useful information, that could not be detailed in the text and help to understand context or other aspects presented				
8.3. All annexes are referenced in the text and are included in the Annexes section, in the order they are referenced				
8.4. Data and information in the annexes are clearly presented and integrate the text				
8.5. Comments:				
9. General considerations				
9.1. The report is written clearly and set out logically				
9.2. The report presents a point of view independent and is not influenced by any stakeholder				
9.3. Specialized concepts are used only when necessary and clearly described (when useful, a glossary is included)				
9.4. Cross-cutting issues such as: (i) gender; (ii) tripartite and social dialogue issues (iii) international labor standards, (iv) environmental sustainability and (v) medium and long-term effects of capacity development action are assessed (if requested)				
9.5. All data is disaggregated by sex, age, ethnic group, or other relevant demographic categories, where feasible;				
9.6. Charts, tables, and graphs are understandable and appropriately and consistently labeled				
9.7. The report addresses the demand of the commissioner/s and is useful				
9.8. Comments:				

REFERENCES

European Commission, DG Regio, 2013 EVALSED :The resource for the evaluation of Socio-Economic Development - Evaluation guide; https://ec.europa.eu/regional_policy/en/information/publications/evaluations-guidance-documents/2013/evalsed-the-resource-for-the-evaluation-of-socio-economic-development-evaluation-guide

European Commission, DG Regio, 2015, Public Procurement - Guidance for practitioners on the avoidance of the most common errors in projects funded by the European Structural and Investment Funds; https://ec.europa.eu/regional_policy/sources/docgener/informat/2014/guidance_public_proc_en.pdf

European Commission, DG Employment, 2021, Design and commissioning of counterfactual impact evaluations - Practical guide; https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8426&further Pubs=yes

Knowlton L. and Phillips, C., 2013, *The Logic Model Guidebook*: Better Strategies for Great Results, SAGE-

ILO, 2014, Checklist 4.2: Preparing the evaluation report, https://www.ilo.org/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_746808.pdf

Miron, G. 2004, Evaluation Report Checklist, Western Michigan University, https://wmich.edu/sites/default/files/attachments/u350/2014/evaluation-reports.pdf

New South Wales Government, Department of Premier and Cabinet Evaluation Toolkit, https://www.dpc.nsw.gov.au/tools-and-resources/evaluation-toolkit/.

Stern, E. 2015, *Impact Evaluation: A Guide for Commissioners and Managers*. Prepared for the Big Lottery Fund, Bond, Comic Relief and the UK Department for International Development.

Wingate L. and Schroeter D., 2016, *Evaluation Questions Checklist for Program Evaluation*, Western Michigan University. https://wmich.edu/evaluation/checklists

RESOURCES

European Commission, 2021, Staff Working Document, Performance, monitoring and evaluation of the European Regional Development Fund, the Cohesion Fund and the Just Transition Fund in 2021-2027; https://ec.europa.eu/regional_policy/en/information/publications/evaluations-guidance-documents/2021/performance-monitoring-and-evaluation-of-the-european-regional-development-fund-the-cohesion-fund-and-the-just-transition-fund-in-2021-2027

European Commission, 2021, "Better Regulation" Guidelines and Toolbox, https://ec.europa.eu/info/law/law-making-process/planning-and-proposing-law/better-regulation-why-and-how/better-regulation-guidelines-and-toolbox_en

Funnell, S.C. and Rogers, P.J., 2011, Purposeful Program Theory: Effective Use of Theories of Change and Logic Models, Ed. Jossey-Bass,

Leeuw F. Vaessen J., 2009, Impact Evaluations and Development, Nonie Guidance on Impact Evaluation, NONIE—The Network of Networks on Impact Evaluation, https://documents1.worldbank.org/curated/en/411821468313779505/pdf/574900WP0Box351BLIC10nonie1guidance.pdf

OECD, 2002, Glossary of Key Terms in Evaluation and Results Based Management, https://www.oecd.org/dac/evaluation/2754804.pdf

Rogers, P. 2014. Theory of Change, Methodological Briefs: Impact Evaluation 2, UNICEF Office of Research, Florence; https://www.unicef-irc.org/publications/747-theory-of-change-methodological-briefs-impact-evaluation-no-2.html

UNICEF, 2016, Series "Methodological Briefs", https://www.unicef-irc.org/publications/series/methodological-briefs/

World Bank, 2011, Writing Terms of Reference for An Evaluation: A How-To Guide, https://ec.europa.eu/regional_policy/sources/docgener/evaluation/doc/05102016/prep_writing_tor.pdf

World Bank, 2015, Managing Evaluations: A How-To Guide For Managers and Commissioners of Evaluations, https://ieg.worldbankgroup.org/sites/default/files/Data/reports/ecd_man_evals.pdf

Useful web sites:

https://www.betterevaluation.org/

https://ieg.worldbankgroup.org/

https://wmich.edu/evaluation/checklists