Evaluation handbook
for EU decentralised agencies

EUAN Performance
Development
Network

2025

EU Agencies Network

agencies-network.europa.eu

Current members and observers of the EUAN	
ACER – member	Agency for the Cooperation of Energy Regulators
AMLA – observer	The Authority for Anti-Money Laundering and Countering the Financing of Terrorism
BEREC Office – member	Agency for Support for BEREC (the BEREC Office)
CBE JU – member	Circular Bio-based Europe Joint Undertaking
CdT – member	Translation Centre for the Bodies of the European Union
CEDEFOP – member	European Centre for the Development of Vocational Training
CEPOL – member	The European Union Agency for Law Enforcement Training
Chips JU – member	Chips Joint Undertaking
Clean Aviation JU – member	Clean Aviation Joint Undertaking
Clean Hydrogen JU – member	Clean Hydrogen Joint Undertaking
CPVO – member	Community Plant Variety Office
EASA – member	European Aviation Safety Agency
EBA – member	European Banking Authority
ECCC – member	European Cybersecurity Competence Centre and Network
ECDC – member	European Centre for Disease Prevention and Control
ECHA – member	European Chemicals Agency
EDA – observer	European Defence Agency
EEA – member	European Environment Agency
EFCA – member	European Fisheries Control Agency
EFSA – member	European Food Safety Authority
EIGE – member	European Institute for Gender Equality
EIOPA – member	European Insurance and Occupational Pensions Authority
EIT – member	European Institute of Innovation & Technology
ELA – member	European Labour Authority
EMA – member	European Medicines Agency
EMSA – member	European Maritime Safety Agency
ENISA – member	European Union Agency for Cybersecurity
EPPO – observer	The European Public Prosecutor's Office
ERA – member	European Union Agency for Railways
ESMA – member	European Securities and Markets Authority
ETF – member	European Training Foundation
EUAA – member	European Union Agency for Asylum

EUDA – member	European Union Drugs Agency
EUIPO – member	European Union Intellectual Property Office
EUISS – member	European Union Institute for Security Studies
eu-LISA – member	European Union Agency for the Operational Management of Large-Scale IT Systems in the Area of Freedom, Security and Justice
EU-OSHA – member	European Agency for Safety and Health at Work
EU-Rail – member	Europe's Rail Joint Undertaking
EUROFOUND – member	European Foundation for the Improvement of Living and Working Conditions
EuroHPC – member	European High-Performance Computing Joint Undertaking
EUROJUST – member	European Union Agency for Criminal Justice Cooperation
EUROPOL – member	European Agency for Law Enforcement Cooperation
EUSPA – member	European Union Agency for the Space Programme
F4E – member	European Joint Undertaking for ITER and the Development of Fusion Energy
FRA – member	European Union Agency for Fundamental Rights
FRONTEX – member	European Border and Coast Guard Agency
Global Health EDCTP3	The Global Health EDCTP3 Joint Undertaking
IHI – member	Innovative Health Initiative Joint Undertaking
SatCen – member	European Union Satellite Centre
SESAR 3 JU – member	SESAR Joint 3 Undertaking
SNS JU	Smart Networks & Services Joint Undertaking
SRB – member	Single Resolution Board

Table of Contents

Preamble

Evaluations have played an increasingly important role in EU decentralised agencies over time. In recent years, this role has been consolidated. Assessing and documenting the effectiveness, efficiency, coherence, relevance, and EU added value of agencies' work is now a key priority for ensuring accountability and organisational learning.

A robust evaluation culture is central to the mission of EU agencies, enabling them to fulfil their mandates effectively while maintaining public trust. Evaluations are not merely tools for assessing past performance; they are mechanisms for shaping the future through informed decision-making. By promoting a culture of continuous improvement, agencies can better align their actions with overarching EU priorities and stakeholder expectations.

The minimum requirements for the agencies' evaluation work are those established in the Framework Financial Regulation[1]. These requirements may sometimes be complemented by requirements in the agencies' founding regulation. Additional helpful guidance and tools can be found in the Commission's Better Regulation guidelines and toolbox[2]. It is important to bear in mind, however, that the Better Regulation guidelines and toolbox were developed for Commission services and in case of doubt, it is recommended to refer to the Framework Financial Regulation.

This handbook aims to provide a clear understanding of how coherent evaluation practices can be established and maintained across EU agencies. Supported by the EUAN PDN Evaluation Working Group's catalogue of templates and good practices, it equips agencies with the tools to adapt the general evaluation principles proportionally to their individual contexts, while ensuring quality and accountability standards.

Scope and structure of the handbook

The first draft of the handbook was circulated in 2014, with an update made in 2017. This 2025 update takes into account developments since 2017 and the experience the agencies have accumulated since then.

This update introduces key concepts and issues regarding evaluation in EU agencies and complements a wider collection of available online resources (good practices, templates, etc.)[3]. The aim is not to reproduce the existing material, but rather to direct the reader towards it.

The handbook should be complemented by additional resources, including sector-specific guidelines, academic research, and peer collaboration. Users are encouraged to remain informed about evolving trends, such as artificial intelligence, innovative research methodologies, and evolving EU policies and regulations. As evaluation tools and technologies advance, EU agencies must adapt their processes to remain effective and relevant. Continuous learning and engagement with emerging practices, including those shared by the EUAN PDN Evaluation Working Group, will ensure that evaluations are robust, forward-looking, and aligned with the evolving mandates of EU agencies.

An increasing number of agencies have provisions in their founding regulations which foresee the European Commission conducting regular evaluations of the agencies' performance. This document does not cover these evaluations. It should be mentioned, however, that for such evaluations, the Better Regulation guidelines apply.

This handbook highlights key aspects of evaluations, including governance, methodological rigour, and effective follow-up of recommendations. It is structured to guide EU agencies through each stage of the evaluation process.

- Chapter 1 introduces the purpose and principles of evaluations, distinguishing between ex-ante and ex-post evaluations.
- Chapter 2 outlines governance frameworks, including planning, coordination, and quality assurance mechanisms that underpin successful evaluations.
- Chapter 3 focuses on the practical aspects of managing evaluations, from design to dissemination, while providing clear methodologies for data collection and analysis.

---

1 Commission Delegated Regulation (EU) 2019/715 of 18 December 2018 on the framework financial regulation for the bodies set up under the TFEU and Euratom Treaty and referred to in Article 70 of Regulation (EU, Euratom) 2018/1046 of the European Parliament and of the Council, C/2018/8599.

2 https://commission.europa.eu/law/law-making-process/planning-and-proposing-law/better-regulation/better-regulation-guidelines-and-toolbox_en

3 To be developed during 2025/2026.

- Annex I provides a checklist for EU agency evaluations.
- Annex II introduces a selection of the most relevant evaluation designs and approaches for agencies.
- Annex III includes examples of evaluation questions aligned with the standard evaluation criteria.
- Annex IV presents examples of evaluation matrices.

EU agencies operate in diverse contexts, ranging from environmental monitoring to humanitarian assistance. This diversity necessitates flexible application of the principles outlined in the handbook, and it is clear that no single evaluation model will fit the needs of all agencies. Evaluation methodologies must account for the complexity, scale, and resources of each agency, ensuring proportionality in their application.

This handbook was prepared by a project group under the PDN Evaluation Working Group:

- Jesper Bejer, EU-OSHA
- Soo Mee Baumann, FRA
- Barbara Floretti, EFSA
- Ilaria Piccioli, EU-OSHA
- Koen Rossel-Cambier, EUAA
- Maria Fatas Ortega, EUAA
- Tine Soerensen Ballis, Cedefop
- EUIPO
- Roxana Iordache, EIOPA

The drafting of the handbook was finalised on 12 May 2025, and it was first endorsed at the Performance Development Network meeting on 13 June 2025 and later at the EUAN Heads of Resources meeting on 22 October 2025.

1 Evaluations in the context of EU agencies

1.1 What is an 'evaluation'

The Better Regulation guidelines define an **'evaluation'** as an evidence-based assessment of the extent to which an intervention is, or is likely to be:

- **effective** in fulfilling expectations and meeting its objectives;
- **efficient** in terms of cost-effectiveness and proportionality of actual costs to benefits;
- **relevant** to current and emerging needs;
- **coherent**;
- of **EU added value**.

In the context of EU agencies' activities, evaluations are systematic, forward-looking or reflective, and evidence-based judgements of a planned, ongoing, or completed intervention (project, programme, or strategy), including its design, implementation, and results.[4]

The systematic and objective assessment is based on a set of criteria and indicators, with the aim of establishing findings and conclusions as well as recommendations for future planning and improvements to the intervention. Evaluations not only identify what has happened, or is likely to happen for planned interventions, but also aim to explain why, and where relevant and feasible, how much of what has happened can be attributed to the intervention being evaluated.

1.1.1 Evaluation and related activities[5]

Evaluation and monitoring: Monitoring is the systematic collection of data during the implementation of an intervention, aimed at tracking progress and enabling timely corrective measures. Monitoring data are an important input to evaluations.

Evaluation and audit: Audit, in a narrow sense, verifies the legality and regularity of the use of inputs against pre-established regulations and standards, and most often aims to detect and prevent irregularities.

1.2 The purpose of evaluation in EU agencies

As public sector organisations, EU agencies are required to demonstrate their performance, achievements, and impact. Amid growing budgetary constraints, there are increasing demands from internal and external stakeholders for good governance, accountability and transparency, greater effectiveness and efficiency, and the delivery of results.

Evaluations aim to inform decision-making by assessing specific interventions and ensuring that relevant evidence is available to support their initiation (the 'evaluate first' principle), to assess them during and/or at the end of their implementation, as well as to inform the preparation of new or follow-up initiatives.

The Better Regulation guidelines indicate that evaluations serve several purposes. Although their importance may differ, most evaluation results will contribute to:

- **timely and relevant advice to decision-making and priority-setting**: evaluations support strategic planning and the design of future interventions, boosting their relevance, effectiveness, efficiency, coherence, and added value;

---

[4] The term 'intervention' will be used throughout this document to refer to the subject of an evaluation.
[5] See also the European Commission's Better Regulation Tool #43.

- **organisational learning**: the results of an evaluation can improve the quality of interventions and contribute to institutional knowledge-building. Evaluations should not only identify areas for improvement, but also encourage the sharing of (good and bad) practices and achievements. Evaluations also provide the opportunity to assess the unintended and/or unexpected effects of the agency's interventions;
- **transparency and accountability**: stakeholders and the general public have the right to know what the agency has done and achieved, and how this has been implemented;
- **efficient resource allocation**: evaluation results support a more efficient allocation of resources between and within interventions. Evaluations can be a useful tool for identifying potential savings, or for achieving a bigger impact with the same level of resources. Across interventions, evaluations can provide evidence on how to best prioritise resources and become more agile. Where relevant, evaluations can also explore opportunities for generating revenue.

While the object of an evaluation is often a specific intervention, it can also be a tool used across various interventions or specific themes covered by several interventions.

EU agencies should promote an evaluation culture that enables internal and external stakeholders to reflect on lessons learnt and best practices. In this context, evaluations include recommendations, which should be discussed and followed up collaboratively to achieve the intended benefits.

1.3 Principles

A set of core principles[6] should be respected when preparing or implementing evaluations:

Evaluate first: The 'evaluate first' principle ensures that decision-making takes due account of lessons learnt from past actions, as shown in the figure below.

Figure: 'Evaluate first' principle
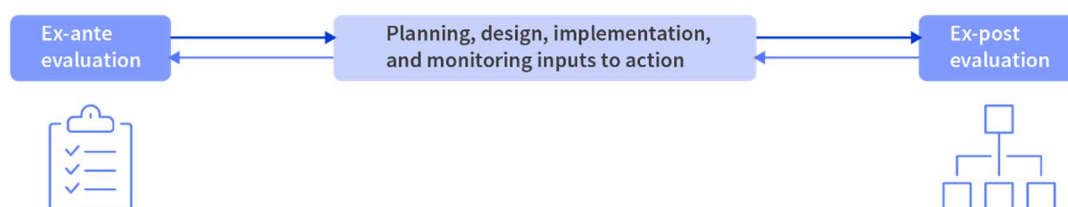
Objectivity and independence: Evaluators must not be in a conflict of interest regarding the intervention and must be free to carry out the evaluation according to professional standards.[7]

Evidence-based reasoning: Evaluations must draw on various, reliable data sources and, where possible, analytical methods.

Proportionality: The scope, data collection, and analysis must be appropriate to the purpose of the evaluation, the nature of the intervention, and practical issues such as data availability and resources allocated to the evaluation.

Comprehensiveness: Evaluation criteria should align with the objectives of the evaluation and should typically include effectiveness, efficiency, relevance, coherence, and EU added value.

Transparency: Evaluations must be based on available evidence and analysis. The criteria used to make judgements should be clearly identified during the evaluation design (see 'Inception phase' below). Any limitations related to methodology and evidence must be reported.

---

[6] See also the European Commission's Better Regulation guidelines, SWD(2021) 305 final, pp. 26-27.
[7] See also the European Commission's Better Regulation Tool #45.

1.4 Ex-ante and ex-post evaluations

The Framework Financial Regulation distinguishes between two types of evaluations based on when they take place: ex-ante evaluations, taking place before the implementation of an intervention, and ex-post evaluations, taking place after an intervention (or after a significant milestone).

In addition to this distinction, a number of evaluation approaches and designs can be identified. A selection of these considered relevant for agencies is presented in Annex II.

Ex-ante evaluations

ex-ante evaluation supports decision-making on future interventions and their design. As stated in the Framework Financial Regulation, interventions involving significant spending should be subject to ex-ante evaluations. These evaluations should be proportionate to the objectives and anticipated costs. However, the Framework Financial Regulation does not specify a precise threshold of what constitutes 'significant spending'; therefore, it is up to each EU agency to define this.

Although the Better Regulation guidelines do not fully apply to the evaluations carried out by the agencies, the toolbox provides useful guidance – see Chapter 2 on impact assessments, that are the Better Regulation equivalent to ex-ante evaluations.

In accordance with Article 29(2) of the Framework Financial Regulation:

> An ex-ante evaluation supporting the preparation of programmes and activities shall be based on evidence, if available, on the performance of related programmes or activities and shall identify and analyse the issues to be addressed, the added value of Union involvement, objectives, expected effects of different options and monitoring and evaluation arrangements.

In the EU agency context, an ex-ante evaluation should identify and analyse the following core elements (as part of the intervention logic):

1. **The issues to be addressed[8] (needs)**
   What is the problem and why is it a problem? What are the consequences? What are the drivers or causes of the problem? How would a causal problem tree look like?

2. **The added value of the EU's involvement**
   Why should the EU (and the EU agency) act? What type of support can the EU agency provide that is not available at Member State level? How would the EU agency's involvement complement other planned or ongoing interventions? How does the intervention align with EU policy priorities?

3. **The objectives**
   What should be achieved? Could objectives be described in a causal results chain? How do they address the causes of the problem identified? Are they SMART[9]?

4. **The expected effects of the different options**
   What options are available to achieve the objectives? What are their expected effects? How do the options compare, e.g. in terms of efficiency or coherence?

5. **Monitoring and evaluation arrangements**
   How will the implementation of the intervention be monitored? How will future evaluations be organised? Considering the intervention logic, what are the priority baselines and indicators to be tracked?

The timing of an ex-ante evaluation is important and depends on its purpose. If the purpose is to assess whether to proceed with an intervention, it should be conducted before a decision on initiating the intervention is taken. If the purpose is to provide input to the design of an intervention, the ex-ante evaluation should be conducted well before the intervention is initiated.

---

[8] This should be based, where available, on evidence of the performance of related programmes or activities.
[9] Specific, measurable, achievable, relevant, and time-bound.

As ex-ante evaluations focus on the future, it may be relevant to incorporate elements of strategic foresight[10], especially for large-scale and long-term interventions. For strategic interventions, using the megatrends framework may be useful.[11]

Ex-post evaluations

Ex-post evaluations (also referred to as 'retrospective evaluations') are evaluations taking place after an intervention concludes or at after completion of a defined key milestone. In the latter case, they may be called mid-term evaluations. Where relevant and feasible, ex-post evaluations go beyond an assessment of what has happened; they consider why something has occurred, including identifying any causal links to the intervention, and, if possible, how much has changed as a result of the intervention. Timing is important, particularly when the outcome of the ex-post evaluation is needed to inform other decisions, such as the design of similar or follow-up interventions. In some cases, a compromise must be found between providing timely information for decision-making and identifying results of the intervention.

**In accordance with Article 29(3) of the Framework Financial Regulation:**

> Retrospective evaluations shall assess the performance of the programme or activity, including aspects such as effectiveness, efficiency, coherence, relevance and EU added value. Retrospective evaluations shall be based on the information generated by the monitoring arrangements and indicators established for the action concerned. They shall be undertaken periodically and in sufficient time for the findings to be taken into account in ex-ante evaluations or impact assessments that support the preparation of related programmes and activities.

In the EU agency context, an ex-post evaluation would typically assess the following core criteria:

- **effectiveness:** fulfilling expectations and meeting the intervention's objectives;
- **efficiency:** cost-effectiveness and proportionality of actual costs to benefits;
- **relevance:** to current and emerging needs;
- **coherence:** internal (with other agency interventions) and external (with other actors);
- **EU added value:** the extent to which results exceed what Member States would have achieved independently.

The selection of criteria should be in line with the objectives of the evaluation.

The five evaluation criteria[12]:

---
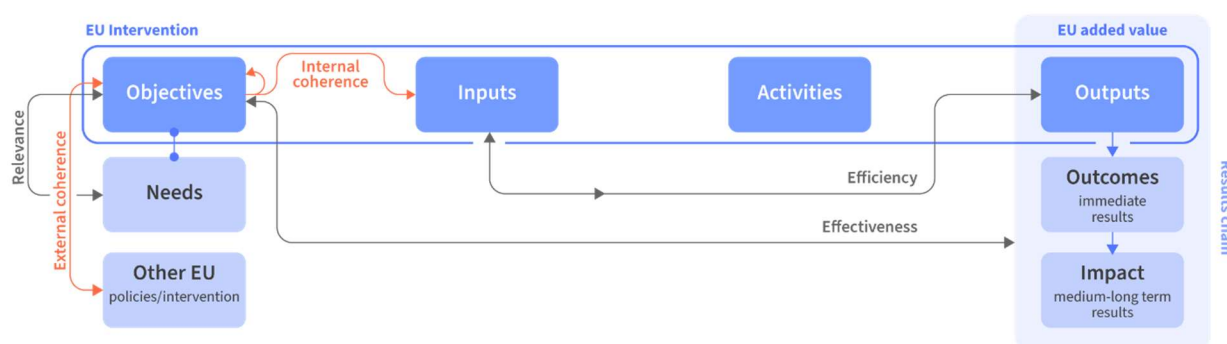
[10] See also the European Commission's Better Regulation Tool #20.

[11] See also the Megatrends Hub

[12] Adapted from the European Commission, Better Regulation Toolbox 2023, Tool #47 available here: https://commission.europa.eu/law/law-making-process/better-regulation/better-regulation-guidelines-and-toolbox/better-regulation-toolbox_en on the basis of Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

2 Evaluation governance at agency level

EU agencies should define and establish minimum governance arrangements to facilitate the commissioning and management of evaluations.

These may include the preparation of an evaluation policy or framework defining the agency's objectives and setting minimum standards, including the follow-up to evaluations. The evaluation policy should also define roles and responsibilities.

The evaluation policy or framework should be complemented by a multi-annual evaluation plan to ensure timely availability of evaluation input and compliance with the Framework Financial Regulation. This plan can also outline the specific mandate, scope, and key stakeholders of planned evaluations.

2.1 EU agency-wide evaluation planning

Evaluations should be planned in a timely and consistent manner to ensure that their findings are available in due time for decision-making. Proper planning improves organisational learning, transparency, and demonstrates the accountability of the EU agency. A first step in conducting evaluations is deciding what to evaluate, which is guided by various criteria.[13]

The Framework Financial Regulation does not automatically subject all programmes and activities to ex-ante or ex-post evaluations. Prioritising and planning evaluations lies within the remit of EU Agencies and should be based on clear criteria and a case-by-case assessment. When evaluations are mandatory due to a legal requirement, which may, for example, be foreseen in the agency's Founding Regulation or specific mandates, they should be considered high priority. Article 29(1) of the Framework Financial Regulation states: 'Programmes and activities that entail significant spending shall be subject to ex-ante and retrospective evaluations ('evaluation'), which shall be proportionate to the objectives and expenditure'. Interventions should be evaluated ex-ante because of their important financial or human resource investment. EU agencies are encouraged to define the financial threshold beyond which an evaluation becomes mandatory. This threshold could ideally be expressed as a percentage of the annual agency budget but could also be expressed in absolute terms. Following the same logic, once completed, these interventions should also be prioritised for ex-post evaluations.

Beyond financial considerations, there may be other situations that necessitate an evaluation. For example, in the case of ex-ante evaluations, an intervention of a novel nature may introduce a significant change requiring senior management or management board review. While such an intervention may not be significant in terms of budgetary resources, it can be critical to the mission and the objectives of the agency, and therefore entail important internal and external risks or commitments (e.g. legal, reputational, financial). In the case of ex-post evaluations, some EU agency strategic documents, programmes, contracts or projects may explicitly anticipate the use of an evaluation – for example, when it is requested by a governing body or the management of the EU agency, when there is a strategic potential to draw lessons which can be used when replicating or expanding an intervention, or when there is a critical need to improve a frequently used working process or service.

Evaluations are not required in all circumstances, and there are scenarios where other approaches are more appropriate. Alternatives such as auditing and monitoring may be more suitable in some cases.

Evaluations should be proportionate, timely, and useful. They are generally not applicable when:

- the situation is covered by other processes, is of an urgent nature and/or has no major impact on existing workflows, e.g.:

---

[13] Please also refer to the European Commission: Directorate-General for International Partnerships, Hassnain, H., McHugh, K., Lorenzoni, M., Alvarez, V. et al., "Evaluation handbook – 2025 edition", Publications Office of the European Union, 2025, Chapter 3.1 on conducting an evaluability assessment.

- the intervention is duly justified as urgent, and the delays linked to a comprehensive ex-ante evaluation would disproportionally affect the expected results;
- the intervention is already evaluated by other entities (partner countries, the European Commission, etc.).
- the cost of the evaluation is disproportionate to the budget of the intervention.

Once the evaluation priorities are defined, the next step is to plan the evaluations. EU agencies should aim for multi-annual evaluation planning to ensure that the required evaluations are identified in advance, internal and external resources are available or secured, and coordination between centres and units is optimised. All potential areas for evaluation should be included in the evaluation plan, spread as evenly as possible over the planning period, without compromising other objectives, such as providing timely input to decision-making.

The planning of the agencies' evaluation activities can be consolidated in a multi-annual indicative rolling plan[14], which is broadly fixed for the first year. It should also provide an agreed plan for the second year and an indicative preliminary plan for the third year. These three-year plans can be reflected in the single programming documents (hereinafter 'SPDs').

The plan can include actions already identified for evaluation beyond the three-year period (e.g. a provision in a five-year strategy that foresees an evaluation towards the end) or other potential future evaluations. The listed actions should be integrated into the rolling plan as it progresses over time.

The most recent version of the plan should be periodically reviewed with management and can be included in the applicable SPD being drafted or updated.

The implementation of the EU agency evaluation plan should be coordinated and monitored via periodic updates. This can create synergies across evaluations in terms of planning, procurement, implementation, and knowledge sharing, contributing to:

- budget and procurement efficiency gains;
- stronger, evidence-based decision-making and change management via horizontal recommendations based on several evaluations;
- enhanced internal learning, cooperation, and knowledge development;
- reinforcing a culture of transparency and accountability through capacity building, networking, and communication.

2.2 EU agency-wide evaluation coordination by its evaluation function

All EU agency evaluations should be coordinated by an evaluation function, in collaboration with the relevant stakeholders.

Agencies should therefore consider establishing a dedicated, centralised evaluation function, composed of evaluation coordinators responsible for coordinating and harmonising all evaluation-related matters. This function acts as a bridge between the agency's corporate governance levels and the respective evaluations, and should be strategically positioned to ensure credibility and impartiality.

The evaluation function can play an overarching role in aligning agency evaluations with the evaluation standards by providing capacity building, technical assistance, and direct support. It can also support design, validation, dissemination, follow-up, and ensure compliance with external reporting obligations related to evaluation planning and follow-up.

Annex I includes a checklist for the establishment of an evaluation function based on good practices.

---

[14] A rolling evaluation plan is required in accordance with the European Commission's Better Regulation Tool #45

2.3 Public access to EU agency documents and the protection of personal data

The principle of transparency and the right of EU citizens and any natural or legal person residing or having its registered office in a Member State to access documents of EU institutions, bodies and agencies are laid down in Article 15 of the TFEU[15] and Article 42 of the EU Charter of Fundamental Rights[16], and implemented through Regulation (EC) No 1049/2001[17] as well as applicable sectoral legislation.

The Better Regulation Tool #49[18] emphasises the importance of communicating and promoting the active use of evaluations and their findings to the widest audience possible, and states that the final evaluation reports should be published in a manner compliant with the agencies' applicable guidelines. Therefore, in principle[19], and for reasons of transparency and openness, it is recommended that the evaluation results be proactively communicated to the public. In its horizontal discharge resolutions on EU agencies, the European Parliament has also repeatedly encouraged EU agencies to share performance information with citizens.

Evaluation results can be disseminated in various formats, not necessarily through publishing a full, lengthy, and technical evaluation report. Alternative channels, such as PowerPoints and infographics, can be used to ensure that the information provided is accessible and understandable by a wider audience.

Depending on the scope, subject matter, purpose of the evaluation, and the stakeholders involved, EU agencies should develop a tailored dissemination plan. The extent and format in which the evaluation findings and recommendations are communicated to the public – beyond the information provided in the Consolidated Annual Activity Report – should consider their usefulness and relevance for the target audiences, the principles of openness and transparency, and the need to cater for organisational learning and improvements.

Before publishing externally, agencies should ensure that there are no justified reasons for withholding publication, in accordance with the exceptions set out in Article 4 of Regulation (EC) No 1049/2001.

Data protection

All publications or disclosures of evaluation deliverables must comply with the applicable data protection rules, including Regulation (EU) 2018/1725. Every effort should be made to avoid including personal data in evaluation reports, both in the main text and annexes, given that reports may be made public. The EU agency's Data Protection Officer (DPO) should be consulted, where necessary.

Legal disclaimer and copyright

All external and internal evaluation reports should include a legal disclaimer and copyright notice. For example, the first page of the report could state:

- 'The sole responsibility of this publication lies with the author. The EU agency is not responsible for any use that may be made of the information contained therein.'
- 'Copyright © for the entire content of this publication unless otherwise stated: xxxxx Agency of the European Union. Re-use is authorised, provided that xxxxx is acknowledged as the source of the material.'

---

[15] Article 15 of the TFEU states: 'Any citizen of the Union, and any natural or legal person residing or having its registered office in a Member State, shall have a right of access to documents of the Union's institutions, bodies, offices and agencies, whatever their medium, subject to the principles and the conditions to be defined in accordance with this paragraph.'

[16] Article 42 of the EU Charter of Fundamental Rights states: 'Any citizen of the Union, and any natural or legal person residing or having its registered office in a Member State, has a right of access to documents of the institutions, bodies, offices and agencies of the Union, whatever their medium.'

[17] Regulation (EC) No 1049/2001 of the European Parliament and of the Council of 30 May 2001 regarding public access to European Parliament, Council and Commission documents (OJ L 145, 31.5.2001, pp. 43–48).

[18] See also the European Commission's Better Regulation Tool #49.

[19] Without prejudice, however, to any other applicable legal provisions, including, inter alia, Article 4 of Regulation (EC) No 1049/2001.

2.4 Evaluation follow-up and reporting

Evaluation reports include recommendations on areas of strength and improvement, addressed to the agency responsible for the implementation of the evaluated intervention. The analysis of the recommendations by the agency, along with the preparation of follow-up actions ensures that the evaluation is acted upon. This analysis should also include an assessment of whether each recommendation is accepted or not. A centralised EU agency evaluation function can enhance this process by supporting and monitoring the follow-up of evaluations.

Article 29(4) and (5) of the Framework Financial Regulation state:

> (4) The director shall prepare an action plan to follow-up on the conclusions of the evaluations referred to in paragraph 3 and report on its progress to the Commission in the consolidated annual activity report referred to in Article 48 and regularly to the management board.
>
> (5) The management board shall scrutinise the implementation of the action plan referred to in paragraph 4.

The management responsible for the evaluation response should ensure that the recommendations and proposed activities remain relevant and applicable. Follow-up actions should be revisited regularly until fully completed.

The evaluation follow-up and reporting help ensure that evaluation findings are fully utilised, hence contributing to greater accountability, good governance, and enhanced organisational learning.

Reporting on follow-up actions should be aligned with the deadlines for the EU agency's legal reporting obligations, such as the Consolidated Annual Activity Report, and be communicated to the EU agency's governance bodies and stakeholders, including the management board.

3 Managing and conducting evaluations

This chapter outlines, firstly, who is involved in an evaluation (the stakeholders) and the various expectations and roles they may have. Secondly, it presents the different phases an evaluation goes through, including the tasks performed at each phase. A clear distinction should be made between the (ex-ante or ex-post) evaluation and the follow-up. Different stakeholders may, for example, be responsible for approving an ex-ante evaluation and for deciding on the follow-up actions based on the evaluation findings.

3.1 Stakeholders and key evaluation actors

Evaluations should ideally involve all relevant internal stakeholders. They are also likely to involve external stakeholders (e.g. contracted evaluation team) and agency stakeholders (e.g. interested parties in the intervention being evaluated). Defining the objectives of an evaluation and its target groups must be done at an early stage, as different target groups have different information needs as indicated below.

- EU Institutional stakeholders (e.g. European Parliament, Council, and Commission) need information on whether the intervention had the desired impact and whether the resources were well spent.
- management boards of EU agencies typically need information to decide on whether to initiate, continue, or repeat an intervention. In this case, the evaluation should aim at clear, concise recommendations.
- managers in charge of the intervention require more detailed information that guides decisions on the intervention's continuation, scale, implementation mode, resource allocation, or other adjustments. Where relevant, they may also require transferable conclusions and recommendations that can be applied to similar interventions.

The objectives of the evaluation should consider the target groups and their needs.

Of utmost importance to the evaluation's success is the timely assignment of the roles and responsibilities, including clarifications on who is responsible for the preparation of the evaluation, its inception, implementation, validation, dissemination, and follow-up. This should consider the actors leading, participating, and contributing to the process, as well as those whose expertise and insights may be required at a given point.

Evaluation manager

The evaluation manager is responsible for coordinating and managing the entire evaluation exercise, from beginning to end. The evaluation manager, typically an EU agency staff member, has the following responsibilities:

- leading the process during the preparation, validation, and dissemination phases, and setting up a quality review mechanism to ensure minimum standards throughout the evaluation;
- overseeing the process and ensuring adherence to quality requirements during the inception, implementation, and follow-up phases (e.g. for work carried out by external contractors);
- ensuring adherence to evaluation standards and a robust methodology, using applicable quality checklists, where relevant.

Evaluation team

The evaluation team conducts the evaluation and communicates closely with the evaluation manager. The team may be comprised of external contractors and/or experts, internal agency staff, or a combination of both and is responsible for the implementation of the evaluation, i.e. data collection, analysis, and report drafting. An external evaluation team should have an appointed team leader. Key responsibilities include:

- delivering the inception and implementation of the evaluation, collecting and analysing data and information, and drafting the inception and evaluation reports;
- ensuring that reports undergo a detailed quality check before submission to the evaluation manager – for external evaluation teams, experienced quality assessors are involved to verify the quality of reports;

- participating in the validation and, where applicable, in the dissemination of the evaluation results.

Implementing actor

The implementing actor is the entity responsible for the intervention under evaluation. This is often a project team, or an agency function and is supported by the relevant agency management. Typically, the responsibilities include:

- providing input on the objectives, scope, and specific evaluation questions during the preparation phase;
- appointing an evaluation focal point who acts as the main point of contact for the evaluation and coordinates efforts;
- collaborating closely with the evaluation manager and evaluation team during the preparation, inception, implementation, validation, and dissemination phases;
- leading the quality review and follow-up by ensuring appropriate follow-up actions and incorporating evaluation findings into the decision-making process. Final responsibility for follow-up lies with the agency's management.

Evaluation Steering Group (ESG)

The Evaluation Steering Group is an interface between the evaluation manager and the evaluation team. Its composition needs to reflect the diversity of stakeholders with an interest in the evaluation. This often includes agency management and staff, members with evaluation expertise, and members familiar with the intervention, such as the implementing actor. The composition of ESG can be adjusted to the specific needs and circumstances of each evaluation. The ESG has typically the following responsibilities:

- advising and supporting the evaluation, and contributing to assessing its quality;
- overseeing the evaluation process along with its deliverables and carrying out a consistency and quality assessment during the inception, implementation, and validation phases;
- ensuring, in close collaboration with the evaluation manager, that the evaluation is evidence-based and unbiased to the extent possible;
- ensuring that the evaluation team has access to and consults all information sources and documentation on activities undertaken;
- supporting the validation of the intervention logic and evaluation questions at the inception phase;
- reviewing and commenting on the reports produced by the evaluation team;
- contributing to lessons learnt and recommendations during the quality review consultations by providing input to the evaluation manager and team.

Depending on the context of an evaluation, additional stakeholders may be involved. These can be individuals, groups, or organisations that have responsibilities and/or direct or indirect interest in the evaluation.

Ethical values and integrity

Ethical values and integrity should be a priority throughout all phases of an evaluation, ensuring that evaluations are conducted in an independent, professional, and transparent manner.

The evaluation should not reflect the interests of a particular person or group of people, or be unduly influenced by the implementing actor, the team, or stakeholders. The presence of stakeholders in the ESG should not compromise the evaluation process.

The evaluation team leader and members must conduct their tasks independently to maintain their impartiality and rigour in the work performed to meet professional standards. At the same time, the team has the responsibility of interpreting and reporting accurately and truthfully the data, information, and documents used during the evaluation.

The evaluation manager and the evaluation team must always behave professionally and be respectful of any views and opinions expressed during data collection. Where the evaluation is fully or partially outsourced, measures to prevent conflicts of interest should be included in the contractual arrangements.

Any potential conflicts of interest must be reported immediately to the evaluation manager, who may consult with the ESG and relevant managers as deemed appropriate.

You may also refer to the principles set out in Section 1.3.

3.2 The different phases of ex-post evaluations

An evaluation can be divided into distinct phases. This section introduces the following:

- preparation phase;
- inception phase;
- implementation phase;
- quality review and validation phase;
- dissemination and follow-up phase.

i. Preparation phase[20]

The preparation phase lays the foundation for how the evaluation will be conducted. Evaluation preparation and design often build on the evaluation plan (see Chapter 2), which typically outlines the purpose, scope, objectives, target groups, and rationale behind the selected evaluation criteria. Since the scope and objectives of an evaluation determine its design, it is crucial to define them at the initial phase to ensure sound implementation. Equally important is ensuring that they are both proportionate to the scale of the intervention. The evaluation manager should also specify the expected deliverables and how the results will feed into the intervention's life cycle.

The evaluation manager should define the main components of an evaluation in the terms of reference. In addition to the elements above, these should include the terms for the dissemination of final deliverables, the target audiences (e.g. the wider public, internal and selected external stakeholders), and the dissemination mechanisms. In the case of an internal evaluation, the terms of reference can take the form of a concept note.

The product of the preparation phase is the terms of reference or a similar document typically including:

- the purpose, scope, objective, and target groups;
- preliminary evaluation questions linked to the evaluation criteria;
- preliminary data collection methods;
- preliminary analytical methods;
- the main outputs and deliverables;
- the evaluation timeline;
- resources available.

Each evaluation should identify priority evaluation questions, bearing in mind the ultimate purpose of the evaluation. In principle, these questions should focus on the five standard evaluation criteria[21] (core elements): relevance, effectiveness, efficiency, coherence, and EU added value. The inclusion or exclusion, in the case of the standard criteria, of the evaluation criteria should be justified. Additional criteria, such as the sustainability or equity of an action, or criteria on specific areas of work, may also be included depending on the overall scope of the evaluation. Particular attention should be paid to impact in view of its connection to the overall intervention logic.

If available (e.g. from an ex-ante evaluation), the intervention logic should be included; otherwise, it can be reconstructed during the inception phase.

---

[20] More information on the preparation phase can be found in the European Commission's Better Regulation Tool #46.
[21] See also Section 1.3 and Section 3.2(ii) 'Inception phase'.

The intervention logic provides a description, often graphical, of how an intervention is expected to work. It outlines a causal chain from the intervention to the expected impacts. It helps foster a shared understanding of how the effects of an intervention are expected to unfold and identify the assumptions required for those impacts to occur.
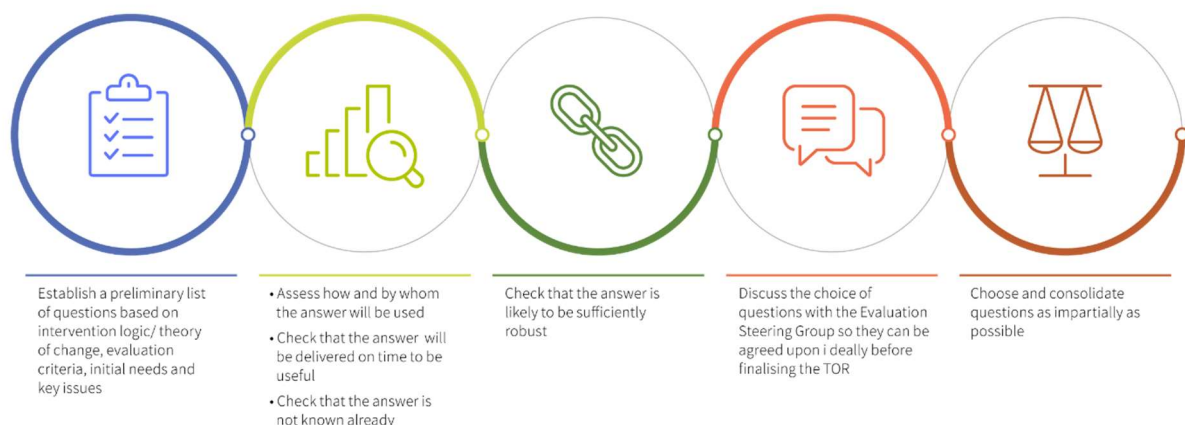
Example of a visualisation of an intervention logic[22]**:**
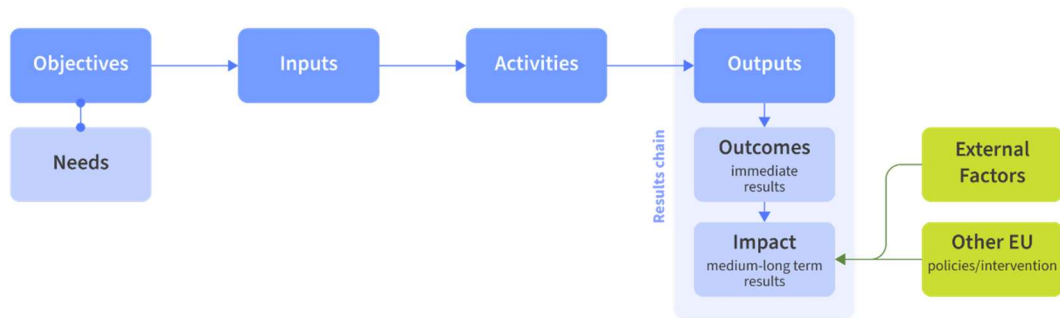
Evaluation questions are a key driver of the evaluation and require particular attention during the preparation phase, as they guide subsequent data collection and analysis. To formulate good evaluation questions:

- use the intervention logic when drafting them;
- ensure questions which stimulate critical analysis and explore links between different elements of the intervention logic;
- use open-ended rather than closed-ended questions;
- complement general questions (e.g. 'To what extent was the intervention effective?') with more specific ones (e.g. 'Which external factors have contributed to or hindered the achievement of the objectives?');
- avoid multiple, redundant, or unanswerable questions;
- avoid prescriptive questions to allow the evaluator room to develop them during the inception phase.

Selecting evaluation questions[23]

Establish a preliminary list of questions based on intervention logic/ theory of change, evaluation criteria, initial needs and key issues

- Assess how and by whom the answer will be used
- Check that the answer will be delivered on time to be useful
- Check that the answer is not known already

Check that the answer is likely to be sufficiently robust

Discuss the choice of questions with the Evaluation Steering Group so they can be agreed upon ideally before finalising the TOR

Choose and consolidate questions as impartially as possible

Examples of ex-post evaluation questions drawing on the five evaluation criteria[24] can be found in Annex III.

The terms of reference are approved by the Evaluation Steering Group (if already established) or the relevant agency management forum.

When the evaluation is carried out externally, the terms of reference serve as the basis for the preparation of the contractor's proposal.

Key roles during the preparation phase

Who	What
Evaluation manager	Prepares terms of reference
Implementing Actor	Provides input to terms of reference
Agency Management or Evaluation Steering Group (if established already)	Approves terms of reference

ii. Inception phase

During the inception phase, the terms of reference should be revisited when the operational arrangements, available evidence, stakeholders, and other factors become available. The EU agencies and evaluators should be flexible in adapting designs to unforeseen risks and emerging issues.

The main aim of the inception phase is that the Evaluation Steering Group and evaluators come to a shared understanding of the evaluation. This will result in an inception report, which typically includes:

- a revised intervention logic;
- revised evaluation questions;
- an evaluation matrix;
- the methodology for data collection, stakeholder consultation, use of case studies etc.;
- identification of methodological risks;
- a detailed work plan.

The data and data collection methods shape the analysis and are essential for effectively addressing the evaluation questions. It is important to consider which data are already available, what additional data may be needed, and consider proportionality issues between data collection tasks and resource availability.

One of the key outputs of the inception phase is the evaluation matrix, which guides data collection and analysis. For each evaluation question, the matrix ideally defines:

- judgement or assessment criteria for answering the evaluation questions;
- indicators and, if possible, targets that inform the judgement;
- data sources for the indicators;
- analytical methods – how data will be turned into indicators.

Examples of evaluation matrices can be found in Annex IV.

---

[24] See also the European Commission's Better Regulation guidelines, Chapter VI

Key roles during the inception phase

Who	What
Evaluation manager	Oversees the inception phase and ensures that minimum quality standards are met
Implementing Actor	Supports the evaluation team and the evaluation manager and provides background information on the intervention and key stakeholders; provides feedback on the (reconstructed) intervention logic and the draft inception report
Evaluation Team	Prepares the inception report
Evaluation Steering Group	Advises and approves the inception report

iii. Implementation phase

The implementation phase focuses on data collection and analysis and includes drafting answers to the evaluation questions using the evaluation matrix.

Data collection methods typically include[25]:

- desk review of documents;
- interviews;
- surveys;
- focus groups.

Data should allow for triangulation, i.e. verification of data from several sources and methods. For analysis, EU agencies generally apply theory-based evaluation approaches (see Annex II), which rely on an intervention logic to establish links between the intervention and its various effects. In most cases, the available data will limit the analytical methods that can be applied.

When reaching out to stakeholders (e.g. through interviews or surveys), the evaluation team must respect personal data protection rules (see Section 2.3) by:

- obtaining prior consent before sharing their personal data with third parties (e.g. contractors);
- informing them on the processing of their personal data during the evaluation, for example, through a privacy statement.

The results of the data collection and analysis are presented in the evaluation report, the main deliverable of the entire evaluation process. The core of the report should be relatively short in length and present the evaluation findings succinctly to facilitate easy access to key recommendations. These recommendations should be supported by key evidence that allows readers to understand how they were derived. An executive summary of the report is also recommended.

The evaluation report should also address the needs of different target audiences. Sections of the final report can be adjusted accordingly – for instance, the opening section of the report can target decision-makers by outlining key findings to inform their decisions.

Evaluation findings should be guided by the intervention logic. The evaluation report should generally cover the following content:

- objectives and scope;
- methodology (including limitations);

---

[25] See also the European Commission's Better Regulation Tool #67.

- expected results of the action (intervention logic);
- analysis of how the situation evolved over the evaluation period;
- evaluation findings;
- conclusions, lessons learnt, and recommendations.

Recommendations should be made in line with the following guidance[26]:

- be action-oriented and realistic;
- clearly stem from the evaluation findings and conclusions;
- identify both the strengths of a given intervention and areas for improvement;
- when appropriate, be directed at specific actors to ensure accountability and adequate response.

Key roles during the implementation phase

Who	What
Implementing actor	Provides required information to evaluation team
Implementing actor	Provides required information to evaluation team
Evaluation team	Implements evaluation and delivers evaluation report
Evaluation Steering Group	Supports the evaluation team with advice and documentation

iv. Quality review and validation phase

A quality review mechanism should ensure that minimum quality standards are met during the evaluation process and across all deliverables.

The evaluation manager, evaluation team, and the ESG are responsible for quality checks at key stages, especially during the inception, implementation, and validation phases.

- the evaluation manager, who coordinates the evaluation process, is ultimately responsible for upholding quality at every stage. Should any issues arise, the manager must inform the ESG.
- the evaluation team is responsible for ensuring that deliverables meet the quality standards (e.g. verifying that the inception report addresses all evaluation questions) before submission to the evaluation manager.
- the ESG oversees the evaluation as a key checkpoint throughout the evaluation. The role of the ESG becomes more prominent during the validation phase, when it reviews all deliverables before their finalisation.

During the validation phase, the evaluation manager and the ESG perform an exhaustive quality assessment of the final report, focusing on several key areas:

- scope of the evaluation (e.g. 'Does the evaluation address the specific questions?')
- contents of the report (e.g. 'Does the report contain all required elements?')
- data (e.g. 'Are the data used in the report accurate and complete?')
- analysis and judgements (e.g. 'Has the methodology been clearly explained and applied consistently?')
- usefulness of recommendations (e.g. 'Have the recommendations been derived from the conclusions?')
- clarity of the report (e.g. 'Is the report logically structured and clearly drafted?')

---

[26] The requirements are adapted from UNICEF's evaluation management response guide. Unicef. Evaluation Office. Version 2.0. January 2018.

Following the quality assessment, a validation meeting should be held with the implementing actor and the relevant stakeholders to review the draft evaluation report before its finalisation by the evaluation team.

Any key actors identified during the preparation phase as relevant to the validation phase (e.g. the implementing actor or specific stakeholders) may be consulted. Their feedback and comments on the draft report should be provided to the evaluation team to be taken into account for the final report.

Following the validation of the evaluation deliverables, the implementing actor (i.e. management) prepares a follow-up action plan in response to the evaluation recommendations. This process offers management the opportunity to respond to the evaluation recommendations, including the possibility of expressing disagreement with specific recommendations.

The management response and action plan may include:

- recommendations made by the evaluation (including any sub-recommendations);
- the degree of acceptance for each recommendation (i.e. not accepted, partially accepted, fully accepted), along with explanations or justifications;
- for partially and fully accepted recommendations:
- improvement action(s) with clearly defined deliverables;
- target dates for completion for each action or set of related actions;
- responsible actor(s) for each action or set of related actions.

The evaluation follow-up action plan serves as the basis for reporting on the status of implementation and follow-up activities.

Key roles during the quality review and validation phase

Who	What
Evaluation manager	Checks report against quality standards
Implementing Actor	Participates in the validation meeting and prepares follow-up action plan
Evaluation Team	Ensures deliverables meet quality standards
Evaluation Steering Group	Reviews deliverables prior to finalisation
Agency management	Adopts the follow-up action plan

v. Dissemination and follow-up phase

Prior to the dissemination of findings, key stakeholders, or by default, their representatives, should be invited to comment on the core recommendations of the evaluation and planned follow-up actions. Once the key messages of the evaluation and the recommendations to be acted upon are agreed (see above), they should be published to ensure transparency and accountability. In addition, this may enhance the impact of the evaluation, for example, by better responding to the information needs of decision-makers and stakeholders.

A good practice is to document the intended communication of results a priori in a communication plan, broken down by different target audiences with an interest in the evaluation. The intended use and dissemination of evaluation findings can also be stipulated a priori in the terms of reference. In view of the above, the terms for dissemination set

out during the preparation phase (and included in the terms of reference) should determine the appropriate channels of communication. These may include more traditional means, such as a final presentation of results and recommendations to management boards, or more creative channels for reaching wider audiences, such as interactive reports, infographics, briefing notes, videos, posts on internal and external newsletters or podcasts, among others.[27]

The evaluation function must regularly monitor the use of the evaluation results. The follow-up on action plans committed to by agency management and/or management boards must also be monitored to ensure the systematic and thorough implementation of evaluation findings and recommendations in accordance with the agreed action plan.

Evaluation results and recommendations are reported in the Consolidated Annual Activity Reports and are used to inform the SPDs, shaping future work and plans. For the SPDs, the Framework Financial Regulation (Article 32(3)(b)) establishes the link between evaluation results and the description of activities in the annual work programme.[28]

Key roles during the dissemination and follow-up phase

Who	What
Evaluation function	Identifies key messages and oversees communication actions. Monitors and reports on the implementation of the follow-up action plan
Implementing actor	Implements improvements as per the follow-up action plan
Agency management	Oversees the implementation of the follow-up action plan

3.3 Conducting ex-ante evaluations

i. Contents

Ex-ante evaluations should be an integral part of the programme development and planning processes in the agencies (i.e. programming, project planning, implementation, and monitoring). A well-designed intervention logic can contribute to more efficient, relevant, and more agile interventions. Ex-ante evaluations can contribute to the development of a baseline and examine the agency's preparedness to undertake specific actions. Typical issues to assess are:

- **legal basis:** the scope of the intervention subject to an ex-ante evaluation should align with the founding regulation of the agency;
- **problem analysis and needs assessment:** the rationale for an intervention is to address specific problems and/or unmet needs in society, public services, or the market. A problem analysis requires an understanding of the causes and consequences of a problem, its scale, and a clear baseline describing how the problem is likely to evolve without the intervention. A needs assessment involves a detailed analysis of the situation, motivations,

---

[27] More information on sharing evaluation findings and recommendations is available at: Disseminating Evaluations | Capacity4dev (europa.eu)

[28] Communication from the Commission on the strengthening of the governance of Union Bodies under Article 70 of the Financial Regulation 2018/1046 and on the guidelines for the Single Programming Document and the Consolidated Annual Activity Report, 20 April 2020, C(2020) 2297 final, Annex 2.

and interests of the key stakeholders. Both exercises require research[29] and stakeholder consultations to identify problems, the needs of stakeholders, and how the agency can best address them;

- **intervention objectives:** an intervention must have clear objectives to define on what basis achievements will be measured. Hence, the objectives of an intervention should be SMART[30], consistent with the agency's strategic and programmatic priorities, and relevant to the identified needs of the target population and problems;

- **intervention options available[31]:** the range of options available for achieving the objectives of the intervention and addressing the causes of the problem identified should be described, where relevant. This includes the identification of interventions (e.g. regulation, co-financing, awareness raising), delivery mechanisms (e.g. direct intervention, support to delivery partners, co-delivery with end-beneficiaries, etc.), as well as associated risks (e.g. failure to achieve objectives, delivery challenges, etc.);

- **option appraisal:** the intervention options should be assessed in terms of their contribution to the objectives and their impacts, and compared both against each other and against the 'do nothing' option. The impacts to be assessed are varied and should – if proportionate – include legal, economic, social, societal, and environmental impacts;

- **preferred option:** based on the option appraisal, the preferred intervention should be described in full. This includes the rationale, including considerations of EU added value, operational objectives, internal and external coherence with agency strategies and EU policy priorities, implementation method, required resources (human and financial), expected results (outputs, outcomes, and impacts), risks and mitigation plans, and lessons learnt from similar interventions in the past;

- **monitoring, evaluation, and reporting**: Monitoring, evaluation, and reporting arrangements should be defined by taking due account of the responsibilities of all government levels involved in the implementation of the intervention. Core indicators for measuring progress towards the intervention's objectives should be laid out, alongside their means of verification and frequency of data collection. The monitoring and evaluation policy of the agency can also inform this process.

ii. The ex-ante evaluation process

Section 3.2 provided an overview of each of the key phases in the evaluation process for ex-post evaluations. Ex-ante evaluations follow largely the same process and have similar roles and responsibilities. During the preparation phase, the scope and deliverables of the evaluation are defined. The inception phase finetunes the different evaluation questions and the intervention logic, while establishing the practical arrangements. The implementation phase builds on the triangulation of different data sources.

The timing of ex-ante evaluations is important. To ensure an objective assessment of different intervention options, ex-ante evaluations should be carried out early in the process. If the ex-ante evaluation is carried out later in the process (e.g. when a preference for a specific option has developed), the objectives of the exercise may have to be adjusted.

The validation and response phase is essential to ensure that all relevant stakeholders provide their input on the preferred option and proposed intervention logic, and to embed a quality review process. The dissemination and follow-up phase also differs from ex-post evaluations, as it can support the launch of a new intervention. Finally, the quality of the data included in the ex-ante evaluation can serve as a baseline for further intervention monitoring and ex-post evaluations.

---

[29] For instance, desk-based research and review of documentation (e.g. work programmes, data from the baseline performance, evaluations, studies, and sources on comparable interventions within or outside the agency), interviews with agency staff responsible for the intervention subject to evaluation, consultations with selected external stakeholders via surveys or interviews, and other relevant sources that may address the problems or unmet needs.

[30] Specific, measurable, achievable, realistic and time bound.

[31] In some cases, only one intervention option is analysed and compared to the 'no intervention' scenario. It is important to ensure that the ex-ante evaluation remains proportionate.

Annex I: Good practice checklist for EU agency evaluation functions

To ensure effective evaluations, EU agencies should consider the following checklist:

Evaluation governance

- Coordination

  - establish a proportionate, dedicated evaluation function to oversee planning, execution, and follow-up;
  - facilitate collaboration among internal and external stakeholders through clearly defined roles and responsibilities;
  - maintain a repository of evaluation outcomes to guide future initiatives.

- Planning

  - develop a multi-annual evaluation plan aligned with the agency's strategic objectives and planning documents (e.g. SPD);
  - prioritise evaluations based on legal obligations, financial significance, synergy opportunities, or strategic importance;
  - ensure that ex-ante evaluations are conducted for significant new initiatives.

- Capacity building

  - provide training and resources for staff involved in evaluation activities;
  - encourage knowledge-sharing across the agencies and beyond (e.g. EUAN Performance Development Network).

Support to EU agency evaluations

- Quality assurance

  - promote the use of steering groups to oversee evaluation design, execution, and validation;
  - ensure that evaluations comply with the principles of comprehensiveness, objectivity and independence, evidence-based reasoning, proportionality, and transparency;
  - review compliance with regulatory requirements, such as those related to personal data protection and confidentiality.

- Data management support

  - promote the design and use of intervention logics and robust results monitoring systems to ensure timely and accurate data collection;
  - advocate for data triangulation methods to enhance the credibility of evaluation findings.

- Stakeholder engagement and communication

  - promote the involvement of internal and external stakeholders in formulating evaluation questions and validating findings;
  - ensure that publication and dissemination strategies are designed to reach diverse audiences in an effective and timely manner.

Evaluation follow-up

- Action plans
    - ensure that recommendations are thoroughly analysed and translated into actionable follow-up plans with clear timelines and responsibilities;
    - monitor implementation progress and report to governance bodies (e.g. via CAAR to MB).

- Feedback loops
    - integrate evaluation learnings into policy and programme design (e.g. updates to SPDs);
    - conduct periodic reviews to assess the long-term impact of implemented recommendations (e.g. through meta-evaluations).

Annex II: Types of evaluations

Annex II introduces various types of evaluations. It is not intended to provide an exhaustive list but rather to indicate what may be relevant to agencies.[32]

Evaluations can be distinguished based on different criteria:

- who carries out the evaluation: internal and external;
- scope: thematic or meta-evaluations;
- focus: results and process;
- methodology: theory-based and impact;
- purpose: formative and summative.

i. Internal and external evaluations

Independence and transparency of the evaluation results are important factors to consider when deciding who will carry out an evaluation. Available resources are another key consideration. Evaluations may be carried out internally, externally, or through a combination of both approaches.

One advantage of internal evaluations is that they foster a sense of ownership and increase acceptance of the feedback and results. Internal evaluations are also valuable for 'formative' evaluations, which support organisational learning and continuous improvement of the evaluated intervention. However, a drawback is that their independence may be questioned. Additionally, internal evaluations may fail to address fundamental issues comprehensively, often recommending incremental improvements rather than deeper reforms. Benchmarking against other organisations – useful in some cases – may also be more difficult to achieve. Some of these concerns can be mitigated through governance arrangements, particularly by setting up an evaluation function (see Chapter 2). Finally, before carrying out an internal evaluation, it is important to ensure the availability of the necessary staff resources.

By contrast, external evaluations are more independent and impartial. External evaluators have sufficient and credible distance from the EU agency staff or work being evaluated. This can lead to more reliable results and, in turn, greater acceptance, especially when subject to stakeholder scrutiny. However, even with external evaluators, it is important to avoid situations of conflict of interests. For example, this could be the case when evaluators have contributed to the implementation or planning of the intervention being evaluated.

ii. Thematic evaluations

While thematic evaluations often focus on one intervention, this is not always the case.

A thematic evaluation examines one or more cross-cutting issues covering several interventions. Examples of themes include gender or working conditions. A variation of this approach is evaluating the use of specific tools used across several interventions, e.g. targeted promotion actions. In thematic evaluations, it is important to clearly define the scope and a limited number of priority evaluation questions to ensure a focused approach.

---

[32] A good overview of the different types of evaluations and relevant considerations can be found in the European Commission: Directorate-General for International Partnerships, Hassnain, H., McHugh, K., Lorenzoni, M., Alvarez, V. et al., "Evaluation handbook – 2025 edition", Publications Office of the European Union, 2025, Section 1.2 'Evaluation types and timing', and Chapter 3 'Approaches, methods and tools'.

iii. Meta-evaluations

In the presence of a significant body of evaluations, it may be useful to identify common conclusions and recommendations across them. This is particularly valuable for organisational learning and can be done through meta-evaluations or evaluation synthesis.

A meta-evaluation can be described as a method for aggregating findings from a number of evaluations. It is based on the premise that a critical analysis of individual evaluations supports organisational learning by identifying patterns, good practices, and areas for improvement. This approach serves several purposes, including:

- promoting a culture of evaluation within the agency by exploring and testing new analytical methods;
- providing a space for cross-cutting evaluation findings to guide change management, including future planning and implementation;
- maximising the impact of existing findings by delving into commonalities, differences, and trends;
- identifying good practices and areas of improvement based on how evaluations have been conducted.

iv. Evaluation of processes and results

Evaluations (typically ex-post) may focus at the same time on the results (or impacts) of a particular intervention ('what') and the procedural aspects of implementation and how these results are achieved ('how'). When focusing on results, the evaluation assesses the achievements of an intervention. When focusing on the process, the evaluation looks at how the activities of an intervention are implemented and how they lead to specific outcomes. A concrete evaluation will often combine elements of both results and process evaluations.

Results evaluations have the following characteristics:

- they develop, collect, and process data across indicators defined in the results framework[33], that is part of the intervention logic to assess the short-, medium-, and long-term effects of an intervention;
- they focus primarily on the effectiveness of an intervention, as their goal is to assess progress towards the intended results.

Process evaluations have the following characteristics:

- they examine what is actually happening in the intervention, including management, staff, and beneficiary behaviour/performance, and how these contribute to the observed results;
- they aim to identify how specific achievements come about and are thus, useful for identifying causal relationships between the components of an intervention and the achievements observed;
- they often focus on efficiency by emphasising the link between the input and output elements of an intervention.

v. Theory-based and impact evaluations

Ex-post evaluations can be broadly categorised into theory-based evaluations and impact evaluations, each employing specific methodologies and serving distinct purposes. Combinations of the two approaches are also possible.

Theory-based evaluations focus on understanding how and why an intervention works. Guided by an intervention logic, they explore the mechanisms linking activities to outcomes. By investigating causal pathways, they help verify if the intervention is unfolding as intended. This approach often employs a combination of qualitative and quantitative methods. For instance, interviews and case studies offer contextual insights, while surveys or monitoring data help validate findings. Triangulating data sources enhances reliability, providing a comprehensive view of the intervention's implementation and its contribution to intended outcomes. While not designed to definitively establish causality,

---

[33] The intervention logic explains how the action's objectives are achieved, including causal relationships and underlying assumptions. Glossary of Key Terms in Evaluation and Results Based Management

theory-based evaluations are essential for assessing programme design, refining strategies, and ensuring alignment with objectives.

Impact evaluations, by contrast, aim to determine whether an intervention achieves its desired effects and the magnitude of those effects. These evaluations are typically underpinned by robust statistical methodologies, such as semi-experimental or experimental designs. Randomised Control Trials (RCTs) are the gold standard for impact evaluations, as they randomly assign participants to treatment and control groups, thereby minimising bias and isolating the intervention's effects. Quasi-experimental designs, such as propensity score matching or natural experiments, offer alternatives when randomisation is not feasible or not suitable for other reasons. These approaches provide credible evidence on causality, making them particularly valuable for scaling effective programmes or informing policy decisions.

While theory-based evaluations delve into the process of achieving outcomes, impact evaluations focus on results. Combining these approaches, where feasible, enriches understanding, capturing not only the efficacy of an intervention but also the reasons behind its success or challenges. This integrated approach yields more nuanced and actionable insights.

vi. Formative and summative evaluations

Ex-post evaluations can also be categorised according to their purpose into formative and summative evaluations. Formative evaluations examine the implementation specificities of an intervention with the aim of improving its delivery. Therefore, their focus is on learning and optimising the management and execution of actions. Summative evaluations, by contrast, focus more on whether an intervention achieved its intended outcomes and to what extent, with a stronger focus on accountability.

Formative evaluations have the following characteristics:

- they are typically carried out during the implementation of an intervention, most commonly in the early phases or at mid-point;
- they gather and process information on the implementation of an intervention (e.g. whether activities are being delivered as expected, which implementation elements seem to be working well or not, or whether adjustments are needed) to support improvements;
- they promote embedded feedback mechanisms that enable staff and stakeholders to share their feedback with evaluators regularly.

Summative evaluations have the following characteristics:

- they are typically carried out after an intervention has concluded, or near its end, with the purpose of assessing whether the anticipated outcomes have been achieved;
- they help generate insights on the value (or long-term impact) of the intervention to help inform future interventions;
- they have high documentation value for the institutional memory by providing information on the achievements of the intervention.

Annex III: Evaluation criteria and evaluation questions

Annex III presents examples of evaluation questions for the five standard evaluation criteria.

Criterion	Typical examples[34]
Relevance	- To what extent did the scope and intended results of the intervention remain relevant over the implementation period? - How did the intended results of the intervention align with broader goals and priorities? - To what extent is the intervention still relevant in view of the intended results? How well do the (original) intended results of the intervention still correspond to current needs? - To what extent do the needs or problems addressed by the intervention continue to require action by the EU agency? - Have the circumstances changed in the meantime so much so that the intervention must adapt accordingly? Will the intervention remain relevant in the foreseeable future? - How have the agency's activities contributed to the achievement of the overall objectives and priorities, and how well have they responded to the needs of the stakeholders?
Effectiveness	- How successful has the EU agency's intervention been in achieving (or progressing towards) its intended results? To what extent were the expected changes delivered? - What have the quantitative and qualitative effects of the intervention been? Do they align with the original expectations? - What external factors have affected progress towards the intended results and how are they linked to the intervention? - If the intended results have not been achieved yet, can they still be achieved in time? If yes, by when? - Have any unexpected or unintended effects driven or hindered progress? What explains these effects?
Efficiency	- What are the costs and benefits for the various stakeholders? Were these distributed as expected? What could explain any differences? - To what extent are the administrative and/or adjustment costs of the intervention justified by the results achieved? - Were the results achieved at the expected cost, or were the costs different? What could explain any differences?

---

[34] Adapted from the European Commission, Better Regulation Toolbox 2023, Tool #47 available here: https://commission.europa.eu/law/law-making-process/better-regulation/better-regulation-guidelines-and-toolbox/better-regulation-toolbox_en on the basis of Creative Commons Attribution 4.0 International (CC BY 4.0) licence

Criterion	Typical examples[34]
	▪ To what extent have factors linked to the intervention affected the efficiency with which the results have been achieved? What other factors influence the costs and benefits? ▪ Have any inefficiencies been identified? Could the intervention have been implemented in a more efficient way? What is the potential for simplification and cost reduction? ▪ If there are significant differences in costs and benefits between Member States, sectors, etc., what is causing them? Were they anticipated? ▪ How timely and efficient is the intervention's administrative process (e.g. reporting and monitoring)? ▪ To what extent has the governance structure of the agency supported its ability to perform its tasks, considering its size, composition, organisation and work processes? To what extent were the activities effectively resourced?
Coherence	▪ To what extent is the intervention coherent with other EU and national actions with similar intended results? ▪ To what extent are the various elements of the intervention coherent with one another? ▪ To what extent have the various elements of the intervention created synergies or balanced possible trade-offs? ▪ To what extent is the intervention aligned with the agency's current priorities? To what extent is the intervention coherent internally? ▪ To what extent is the intervention coherent with international obligations or EU policies? ▪ To what extent has the agency coordinated its work with other relevant bodies, government institutions and agencies at EU, Member State, and international levels carrying out similar tasks? To what extent were overlaps avoided?
EU added value	▪ Are the subsidiarity arguments presented in a prior impact assessment or ex-ante evaluation still valid? ▪ Could the intended results have been sufficiently achieved by the partner organisations acting alone? Would national intervention or the absence of the EU agency's intervention significantly damage the interests of other Member States? ▪ In the absence of the EU agency's intervention, to what extent would Member States or third countries have been able to implement appropriate measures? ▪ If the initial problem and its causes varied across Member States, regional and local levels, did the EU agency's intervention help establish a level playing field? ▪ Were there clear benefits from the EU agency's intervention? ▪ Were there economies of scale or shared services? Were the intended results delivered more efficiently by the EU agency than they would have been by Member States or partner countries acting independently (higher benefits per unit cost)? ▪ Could the identified impacts, outcomes, or outputs have been achieved without the EU agency's intervention?

Criterion	Typical examples[34]
	▪ Is it still valid to assume that the EU agency is best positioned to deliver the intended results of the intervention? ▪ What is the added value of the EU agency's intervention, compared to what could reasonably have been expected from Member States or third countries acting at national and/or regional levels? ▪ What are the most likely consequences of discontinuing or withdrawing the EU agency's existing intervention?

Annex IV: Example and template for evaluation matrices

Evaluation Matrix – Third Foresight Activity (Circular Economy) (2024), example from EU-OSHA

Evaluation questions	Judgement criteria	Indicators	Data sources and analytical approach
Effectiveness			
• What observable outcomes have stemmed so far from the Third Foresight Activity?	• The Third Foresight Activity has generated observable results that comprise all envisioned deliverable/outputs of the activity (e.g. publications, seminars) • It is possible to identify additional results and knock-on/secondary effects stemming from the Third Foresight Activity beyond immediate outputs and deliverables	• Tangible outputs and usable results stemming from the Third Foresight Activity • Evidence of additional results and secondary effects stemming from the Third Foresight Activity (e.g. networking benefits, awareness raising, synergies with Member State-level work)	• Document review • Survey • Focus groups • Interviews

General Evaluation Matrix, example from EUAA

Sub-questions	Indicators / Descriptors	Norms / Judgement criteria	Sources of evidence
Relevance: How well was the action aligned with stakeholders' needs and the agency's objectives?			
Optional prompt question: What activities were most beneficial and why?	▪ Priority areas identified by the ex-ante evaluation	▪ Comparison of the ex-ante evaluation and action priority areas with implemented areas	▪ Needs assessment ▪ Intervention logic ▪ Monitoring and reporting tools ▪ Interviews
Effectiveness: Did the action achieve what was planned?			
Optional prompt question: Were deliverables (workflows and tools) as expected? If not, what are solutions/alternatives to achieve better results?	▪ Results indicators	▪ Comparison planned targets vs achieved results	▪ Results Monitoring Database
Efficiency: To what extent are the costs (including inputs and human resources) of the support justified by the results?			
	▪ Output and input indicator values	▪ Relationship between achieved outputs and inputs ▪ Qualitative challenges (processes and related indicators)	▪ Qualitative and quantitative monitoring data ▪ Financial records ▪ Assessment of processes supporting planned results ▪ Interviews
Coherence: To what extent is the action coherent internally and externally?			
	▪ Nature of activities and coordination processes	▪ Level of coordination and synergies with other national actors/internal agency actors	▪ Monitoring data ▪ Planning documents ▪ Interviews ▪ Mapping from the ex-ante evaluation
EU added value: What is the added value of the action, compared to what could have been expected from the partner organisation acting alone?			
	▪ Existence of EU added value elements	▪ Number and level of added value elements related to financial, technical, and material support	▪ Monitoring data ▪ Planning documents ▪ Interviews