

**OPINION ON POST-2027 ESF+
EVALUATION ARRANGEMENTS PREPARED BY MEMBER STATE EXPERTS
COLLABORATING WITHIN THE ESF+ EVALUATION PARTNERSHIP NETWORK**

1. Background

The current non-paper is informed by discussions among Member State members of the ESF+ Evaluation Partnership and is an expert opinion not binding on the Member States or the European Commission. It outlines effective practices and areas for enhancement in the evaluation process for the post-2027 period. The ESF+ Evaluation Partnership is a collaborative initiative designed to improve the quality and effectiveness of evaluations for programs funded by the European Social Fund Plus (ESF+). This partnership encompasses a diverse group of participants, including national and regional authorities responsible for programme implementation, independent evaluation experts, and representatives from the European Commission. The partnership experts of the Member States should be included in the future process of the design of all evaluation related articles in the CPR of post-2027. The primary objective of the ESF+ Evaluation Partnership is to support evidence-based policymaking by leveraging evaluation results to inform and adjust policies and strategies. Through collective efforts, the ESF+ Evaluation Partnership aims to establish a robust evaluation framework that enhances program outcomes and contributes to the overall success of ESF+ funded initiatives. The partnership aims to elevate evaluation quality by refining methodologies and practices, promote knowledge sharing among stakeholders to exchange best practices and lessons learned, and ensure accountability through transparent, evidence-based assessments of program impact.

The recommendations in this non-paper have been written and approved by the evaluation units of the authorities responsible for the implementation of ESF+ in the member states. They draw on valuable lessons learned from managing evaluations during the current and previous programming periods. Incorporating these lessons into future practices, will support the creation of a more effective and responsive evaluation system that supports continuous improvement and policy refinement.

Evaluation plays a pivotal role in understanding the effectiveness, efficiency, and sustainability of policies and interventions. While audits and monitoring have the significant focus of financial compliance and real-time performance, evaluations provide a structured approach to assessing short-term and long-term outcomes and broader implications of interventions. They offer critical insights into what works, what doesn't, why, and for whom, enabling policymakers to make informed decisions, adjust strategies, and optimize resource allocation. By systematically analysing both successes and shortcomings, evaluations facilitate evidence-based policymaking, ensuring that lessons learned from past initiatives are integrated into future program design and implementation. This continuous feedback loop is essential for driving improvements and achieving sustained positive results, ultimately contributing to more effective and impactful public policies.

To strengthen the evaluation process for the post-2027 period, several key elements need to be enhanced and clarified. These improvements will not only bolster the credibility and utility of evaluations but also provide valuable insights to inform future policy and program development. The paper is organized into two main sections: chapter 2 addressing issues to be included in the Common Provisions Regulation (CPR), and chapter 3 focusing on aspects that need more emphasis in evaluation guidelines. The issues in chapter 2 are meant to be mandatory. The issues in the guidelines chapter are intended as non-binding suggestions and recommendations. You will find that to some extent there is overlap between the issues addressed in the two chapters. In general, in the chapter on guidelines we

have included more aspects and elaborations on the issues raised in chapter 2. In a last section, chapter 4, we have addressed issues for which we would appreciate the support of the European Commission. We have written the paragraphs in the chapters in the order of priority of the Managing Authorities (MAs) / Member States (MS) evaluation unit members of the ESF+ Evaluation Partnership Network.

2. Key recommendations to be addressed in the CPR

Refining the evaluation requirements for EU-funded interventions in the CPR for the post-2027 period is crucial to improving project quality and impact. This process will provide better insights into the effectiveness, sustainability, and long-term benefits of these interventions. Key areas to address include refining evaluation criteria, maintaining flexibility, aligning timing with implementation, ensuring data access, fostering accountability, and building capacity for independent evaluations. These improvements will enhance the ability of EU interventions to deliver lasting benefits for Member States. The issues in the first section are meant to be compulsory.

2.1. Refining the evaluation criteria – replacing EU added value with impact and sustainability

One of the key areas for improvement is the refinement of the evaluation criteria used in assessing EU-funded interventions. Currently the Commission is required to carry out mid-term and ex-post evaluations covering the five Better Regulation evaluation criteria, namely effectiveness, efficiency, relevance, coherence and Union added value. Discussions among evaluation partners have highlighted that the concept of *EU added value* is somewhat vague and is often reduced to questioning whether an intervention would have occurred without EU funding. In this context, the criterion is seen as overly broad and difficult to apply consistently. Therefore, it is recommended that the focus shifts to evaluating *impact* and *sustainability*.* By doing so we align these criteria with [the OECD evaluation standards](#), which could enhance clarity and consistency, as the OECD provides extensive explanations and guidance on effective evaluation practices.

Evaluating the *impact* of interventions involves assessing the tangible changes brought about by the programme in social, economic, or environmental terms. Compared to the criterion EU added value, this approach offers a more concrete and results-driven understanding of how EU-funded projects address the real needs and priorities of beneficiaries and society as a whole. Impact focuses on the extent to which and how the intervention has generated or is expected to generate significant positive or negative effects, whether intended or unintended, as well as higher-level consequences. It addresses the ultimate significance and potentially transformative effects of the intervention, seeking to capture not only the immediate results but also the indirect, secondary, and potential outcomes that may arise from the intervention. *Sustainability*, on the other hand, addresses whether the benefits of a project will endure after the funding period has ended. Projects that are sustainable provide a higher return on investment in the long term, as they continue to generate positive results even in the absence of continued financial support. This criterion allows for a more thorough assessment of whether an intervention has been designed in a way that enables it to persist and deliver value independently of external funding.

Shifting the focus from *Union added value* to *impact* and *sustainability* also offers a stronger methodological foundation for evaluations. This change will enable evaluations to more

accurately reflect the progression of programmes from initial effectiveness to long-term sustainability.

Please also see paragraph 3.3 on clarification of evaluation criteria.

2.2. Flexibility over fixed timelines of evaluations

Traditionally, mid-term and final evaluations have been scheduled according to predefined timelines, as outlined in the CPR. Where the mid-term and ex-post evaluations are carried out by the Commission, a mid-term review and the final evaluations are carried out by the member states. While the strict deadline for the final evaluation is generally seen as fixed, the timing for the mid-term evaluation and review could benefit from more flexibility. One option is to set the mid-term review deadline in relation to the programme's adoption, for example, three years after its start. The decision on how and when a mid-term review will be carried out should remain at MS level.

This flexibility is necessary due to delays in programme adoption and the establishment of new procedures/ set-up in order to comply with new Regulatory requirements and other procedures, which have been significant during the current programming period. Delayed approvals reduce the time available for implementation and evaluation, making the existing fixed timelines less practical. Adjusting the mid-term evaluation timing would ensure a more accurate assessment of the programme's progress and effectiveness, giving sufficient time for meaningful data collection and analysis.

Please also see paragraphs 3.2 and 3.4 on learning from evaluations and flexibility to adjust deadlines.

2.3. Clearer wording on evaluation scope

The CPR could benefit from clearer working on what should be evaluated and what not. For instance, Article 44, paragraph 2 of the current CPR creates ambiguity about whether a full ex-post evaluation of the entire programme is required, or if at least one evaluation of the programme suffices. During 2014-2020 programming period, Member States were strongly encouraged to cover the whole programme. The MAs/MS should determine in their evaluation plan what parts of the programmes should be evaluated. For example, the CPR should clarify whether evaluating the horizontal principles is mandatory or not.

Please also see paragraphs 3.5, 3.6 and 4.1 on the evaluation plan, horizontal principles and the Commission's expectations.

2.4. Ensuring adequate evaluation capacity in the MAs/MS and independence in evaluation processes

Maintain or incorporate sufficient provisions for both the capacity of Managing Authorities / Member States within the CPR could help for enhancing the quality and effectiveness of evaluations for EU-funded interventions. The capacity of MAs/MS to conduct and commission evaluations involves ensuring that they are equipped with the adequate personnel and skills. Evaluations require specialized expertise, including knowledge in data analysis, evaluation methodologies, and programme management.

Addressing the capacity of MAs/MS in the CPR will support ensuring that evaluations are conducted effectively and impartially. Adequate capacity ensures high-quality evaluations, while independence safeguards the objectivity and reliability of findings.

By maintaining the current provisions in the CPR that ensure the independence of evaluation experts, the regulation will continue to promote transparency and accountability, fostering trust among

stakeholders and preserving the credibility of the evaluation process. Furthermore, it ensures that the evaluations remain objective, capable of identifying areas for improvement, and able to provide actionable recommendations that drive meaningful and sustainable programme enhancements.

Please also see paragraphs 3.1, 3.7 and 3.10 on clarification of evaluation, methodologies and independence of evaluation units and experts.

2.5. Legal basis for accessing administrative data for evaluations

This issue has been amply addressed in the expert opinion paper on Monitoring and Reporting. However, since it is key for policy evaluations, it will also be discussed here.

Given the specific realities in each country, the partnership recommends that the issue of access to administrative data be addressed in the CPR. The CPR should give a legal basis for Member States to make administrative data accessible for evaluations, promoting greater transparency and facilitating more robust evaluations, while respecting the right for personal data protection.

Ensuring access to high-quality administrative data is essential for conducting robust and comprehensive evaluations of EU-funded interventions. By enhancing the availability and accessibility of such data, including detailed registry data, MAs/MS can significantly improve the evaluation process and contribute to more effective policymaking. It would allow evaluators to build a more reliable evidence base, leading to accurate and trustworthy findings. This, in turn, supports thorough analysis and assessment, ensuring evaluations are grounded in solid factual information.

Access to a broad range of administrative data, including registries, facilitates more in-depth and comprehensive analyses. Evaluators can explore various dimensions of programme performance, such as demographic, geographic, and sector-specific outcomes. This extensive access supports a deeper understanding of how interventions perform across different contexts and subgroups, leading to more nuanced and actionable insights.

Improved access to administrative data empowers making of well-informed decisions based on empirical evidence. By utilizing this data, a clearer understanding of programme effectiveness is gained, areas needing improvement are identified, and resources are allocated more effectively. This enhanced data access contributes to more targeted and effective policy interventions.

Addressing data gaps and integrating various sources of administrative data, including registries, helps in developing a comprehensive understanding of programme performance. Register data can fill gaps where access to a centralized registry is unavailable, offering reliable insights into key metrics like employment, education, and social services. Improved access to these data sources facilitates cross-referencing and validation, leading to more accurate and reliable evaluation results. Furthermore, combining register data with other methods, such as surveys or interviews, complements the quantitative information by capturing qualitative dimensions like participant experiences. This multi-method approach enables a more holistic and improved understanding of intervention impacts.

Access to administrative registries could enhance the cost-effectiveness of data collection in evaluations. Many Member States currently rely on questionnaires, which can be resource-intensive and costly. Utilizing existing administrative data can help reduce these costs while still ensuring high-quality data. Since registries already contain large-scale, systematically collected information, this approach minimizes the need for additional data collection efforts. Multi-method approaches can then be used in a targeted manner. By promoting access to administrative registries, evaluations can become more efficient and cost-effective, while still providing reliable and timely insights.

In summary, giving a legal basis for Member States to improve access to registry data and other administrative data is crucial for enhancing the evaluation process. By increasing data availability and

streamlining data-sharing procedures, evaluations can become more comprehensive, timely, and relevant. This strengthens the evidence base for policymaking, supports more informed decision-making, and ultimately contributes to more successful programme outcomes and greater value for stakeholders.

Please also see paragraph 3.7 on methodology.

2.6. Keep evaluation plan flexible

According to the CPR for 2021-2027, the Member States are to draw up an evaluation plan for the programme(s). The Evaluation Partnership is of the opinion that the evaluation plan, as a critical component of the evaluation process, should remain a flexible and "living" document capable of adapting to emerging priorities and shifting realities. This adaptability allows MAs/MS to respond dynamically to changes in the socio-political environment or shifts in policy focus, ensuring that evaluations stay relevant, timely, and effective.

A flexible evaluation plan is essential for navigating the dynamic contexts in which EU-funded programmes operate. Rapid and unpredictable changes in political priorities, economic conditions, or unforeseen events, such as economic crises or global pandemics, may require a reassessment of programme objectives and strategies. By maintaining flexibility, the evaluation process can be adjusted in real time, ensuring that the approach remains aligned with new challenges and realities. This ability to adapt is crucial for maintaining the relevance and effectiveness of evaluations, as rigid plans risk becoming disconnected from the programme's evolving goals and broader policy imperatives. A flexible evaluation plan, on the other hand, allows MAs/MS to modify methodologies and timelines to capture the most current and relevant information, providing more actionable insights for decision-making.

A flexible evaluation plan encourages the incorporation of innovative or experimental methodologies. As programmes evolve and new technologies or assessment techniques emerge, evaluation plans should be able to integrate these advancements if needed. This openness to innovation allows for more sophisticated assessments of how new approaches contribute to achieving desired outcomes. Finally, flexibility is key to managing resource constraints. MAs/MS often face limitations in terms of the resources and capacity available for conducting evaluations. A flexible evaluation plan enables the reallocation of resources and the adjustment of evaluation scope when necessary, ensuring that evaluations remain feasible and focused on the most critical aspects of programme performance, even in resource-limited scenarios.

Keeping evaluation plans flexible is essential for adapting to changing circumstances, ensuring relevance, addressing programme-specific needs, encouraging innovation, and managing resource constraints effectively.

Please also see paragraphs 3.4 and 3.5 flexibility to adjust deadlines and evaluation plan.

2.7. Establishing a connection between programming and evaluations

In the programming section of the CPR, it is crucial to emphasize the integration of evaluation considerations into programme planning. Furthermore, anticipating at the programming phase which and how key operations will be evaluated is essential for preparing high-quality evaluations in the future. Establishing evaluation frameworks during this phase enables the identification of relevant indicators, data collection methods, and evaluation criteria from the outset. This foresight ensures that evaluations are not only feasible but also aligned with the programme's objectives, facilitating more meaningful insights and actionable recommendations when the time for evaluation arrives. Ultimately, this proactive approach enhances the overall effectiveness of evaluations and contributes to improved

decision-making and programme design. Member States can use the evaluation frameworks as a basis of their evaluation plan. Like the evaluation plan, the evaluation frameworks should be kept flexible in order to adapt to shifting realities and priorities.

3. Key recommendations to be addressed in the post-2027 evaluation guidelines

This section describes which key recommendations we propose that DG Employment addresses in evaluation guidelines. These guidelines are not to be mandatory for the MAs/MS but act as a toolbox for evaluation post-2027.

3.1. Clarification of evaluation and its distinction from other instruments

Inclusion of a clear explanation of what evaluations are and how they differ from related instruments such as spending reviews, audits, monitoring and other type of studies/analyses in the guidelines is crucial for a comprehensive understanding of their respective roles and benefits. Evaluations are designed to assess the effectiveness, impact, and sustainability of interventions. They focus on understanding what works, what does not, and the reasons behind these outcomes. This approach is distinct from other instruments that serve different purposes.

Spending reviews primarily evaluate the efficiency and cost-effectiveness of expenditures, focusing on whether resources are utilized wisely. Audits are concerned with compliance and accuracy in financial reporting, ensuring that funds are used according to established regulations. Monitoring involves regular oversight of program activities and progress to ensure that interventions are meeting their operational targets.

Clarifying these distinctions within the guidelines helps prevent confusion and ensures that stakeholders understand the unique value of evaluations. Evaluations are not just about assessing individual interventions but also about aggregating findings across multiple interventions to draw broader conclusions. This learning loop is essential for applying lessons from past programs to the design of future initiatives, thereby improving the quality and effectiveness of new policy frameworks. By explaining these differences upfront, policymakers and practitioners can better appreciate the specific contributions of evaluations to the policy process, ensuring that evaluations are effectively utilized to generate actionable insights and drive improvements in future programs.

3.2. Learning from evaluation

To effectively leverage evaluations for influencing future programmes, it is crucial to address the concept of “Learning from Evaluation” in the guidelines. This entails not only employing a diverse range of evaluation methods but also focusing on how aggregate evaluations can inform the design of upcoming programs.

Due to the timing constraints of program evaluations, which often conclude after a program has been fully implemented, leveraging these evaluations to influence the design of upcoming programs can be challenging. It is essential for guidelines to incorporate provisions for evaluations that not only assess current programs for implications for the upcoming programs but also consider their implications for initiatives further in the future. The guidelines should also provide the possibility to plan evaluations that span older programming periods allowing us to draw on more long-term data which will contribute to improving the validity of findings.

To bridge the gap between programme evaluations and future programme design, the guidelines should advocate for evaluations that focus on generating insights relevant to upcoming programmes. This approach ensures that evaluation findings are timely and can be integrated into the planning stages of new interventions, allowing for more responsive and adaptive policy development.

The ability to learn from past evaluations and apply these lessons to future programmes is crucial for continuous improvement. By systematically analysing evaluation results and identifying patterns or trends, policymakers can make informed decisions and refine future programs to address identified challenges and build on successful strategies.

Incorporating these elements into the guidelines will enhance the utility of evaluations not only as a tool for assessing past performance but also as a strategic resource for shaping and optimizing future programmes. By focusing on diverse methodologies and evaluating for future impact, the guidelines will support a more proactive and evidence-based approach to policy development and program design.

3.3. Clarifying the key evaluation criteria

Providing a clear explanation of the core evaluation criteria in guidance documents is crucial for several reasons. In the previous chapter we suggested to align with the OECD criteria (see paragraph 2.1). Key criteria such as relevance, effectiveness, efficiency, coherence, impact, and sustainability must be distinctly defined and explained. These criteria are foundational for conducting rigorous and meaningful evaluations, ensuring that evaluations are both comprehensive and comparable. It could be useful that the guidelines reveal the proper use of evaluation criteria in relation to the progress of programme implementation. This does not mean that we want standardized approaches or evaluation questions across all Member States. The aim is to create support for the planning of evaluation for the Member States. Both the Member States and the European Commission will benefit from this.

Defining these criteria, helps prevent inconsistencies in how evaluations are conducted and interpreted across different programmes and Member States. Conceptualizing these criteria is essential for maintaining evaluation rigor and comparability. By applying these criteria uniformly, stakeholders – including policymakers, evaluators, and practitioners – can accurately assess the success of interventions, ensuring that evaluations provide meaningful and actionable insights. This consistency facilitates an approach that enables the effective synthesis of findings and comparisons across various programs and initiatives.

Moreover, clearly defined criteria offer a structured framework for evaluating to what extent programmes meet their objectives, provide value for money, and contribute to broader policy goals. That will contribute to learning from the findings and to informed policy decision-making. This clarity will enable the European Commission to aggregate and analyse evaluation results effectively, leading to better strategic planning and policy development at both national and EU levels.

Furthermore, detailed explanation of evaluation criteria enhances communication and understanding among stakeholders, since it reduces ambiguity and aligns expectations. This shared understanding is vital for ensuring that evaluations are relevant, actionable, and capable of driving continuous improvement in policy and program design. Including clear definitions of these criteria in evaluation guidelines upholds transparency and accountability in the evaluation process. It ensures that evaluations adhere to established standards, fostering trust in the evaluation findings and reinforcing the commitment to evidence-based policymaking.

3.4. Flexibility allowing to adjust deadlines

The timing of evaluations is a crucial aspect that requires careful consideration (see also paragraphs 2.2 and 2.6). As experience has demonstrated, strict adherence to fixed dates can undermine both the effectiveness and relevance of evaluations and their findings. A more flexible approach, aligned with the actual phase of programme implementation, would allow for more meaningful and timely insights.

Evaluations should ideally be conducted when they reflect the true progress and impact of a programme. Fixed dates may not coincide with important milestones or key phases of implementation, which can result in evaluations that fail to capture the evaluation criteria of the programme accurately. Flexible timing allows evaluations to be scheduled when they can provide the most relevant data, offering a clearer picture of the programme's progress and impact. This adaptability is particularly important in accommodating delays and adjustments that programmes often face, such as those seen in the current programming period. Administrative hurdles, unforeseen challenges, or shifts in priorities can delay programme adoption and implementation. Conducting evaluations during these delays risks producing assessments that are premature or incomplete. By allowing for adjustments to the evaluation schedule, evaluations can be conducted when the programme is fully operational, and its effects can be properly assessed.

Flexible timing also enhances the quality and relevance of evaluation findings. When evaluations are conducted at rigid, predetermined times, the results may either come too early to capture meaningful impacts or too late to influence current decision-making and policy design. Aligning evaluations with periods of significant programme activity or change ensures that the findings are both timely and actionable, directly supporting ongoing programme management and improvement efforts. Additionally, as programmes evolve, new issues and insights may arise that were not anticipated when the initial evaluation schedule was set. Flexible evaluation timing allows these developments to be incorporated into the evaluation process, making the assessments more relevant and effective in addressing current challenges and informing necessary adjustments to the programme.

This flexibility also supports an iterative approach to learning and continuous improvement. It leads to more responsive management and better overall outcomes. Moreover, it ensures that evaluations more accurately reflect real-world conditions. The dynamic nature of programme implementation often clashes with rigid evaluation schedules, resulting in assessments that may not align with the actual progress of the programme. By scheduling evaluations to match real-time developments, the assessments provide more accurate insights and contribute to better-informed decision-making.

In conclusion, allowing flexibility in the timing of evaluations ensures they are conducted at the most relevant points in a programme's lifecycle. It improves the accuracy and applicability of evaluation findings, supports better alignment with programme progress, and enhances overall effectiveness. Moving away from rigid timelines enables evaluations to offer more actionable insights, fostering continuous improvement and more successful programme outcomes.

3.5. What should be incorporated in the evaluation plan

Including a clear explanation of what is relevant to evaluate and what should be incorporated into the evaluation plan of the MA/MS in the guidelines for the post-2027 period is essential for ensuring that evaluations are both effective and resource efficient (see also paragraphs 2.3 and 2.8). For instance, certain low-impact components, such as technical assistance, small-scale projects or projects or priorities that already have been evaluated in the past, may not require in-depth evaluation. This targeted approach will help programme managers and evaluators focus on the most impactful and relevant elements of a programme, avoiding unnecessary or redundant evaluations that can drain resources and lead to confusion. A well-structured evaluation plan allows for prioritization, ensuring that critical areas such as long-term impact, sustainability, and efficiency are properly assessed, while less significant aspects that may not require in-depth evaluation can be identified early on.

By clarifying what is relevant to evaluate, the guidelines will ensure that evaluations are designed to produce actionable insights, improving both policy and program implementation. This section of the guidelines would for example clarify to what extent inclusion of non-EU funded activities could be incorporated in evaluations and financed by ESF+.

These explanations will also enhance comparability across Member States, allowing for better synthesis of findings at the EU level and enabling more informed, evidence-based decision-making. Moreover, by specifying what should be included in the evaluation plan, the guidelines will foster a more strategic approach to evaluations, ensuring that they are aligned with the overall objectives of the program and the broader policy framework.

3.6. Identify the relevant horizontal principles for evaluation and determine how to effectively assess them

The guidelines should include a clearer treatment of horizontal principles providing insights into which aspects are relevant for evaluation and how to effectively assess them. The guidelines can help Member States identify key indicators and appropriate methodologies for evaluating these principles, fostering a deeper understanding of their impact on policy outcomes. Please see also paragraphs 2.3 and 4.1.

3.7. Encouraging the use of diverse methodologies

Expanding the use of evaluation methodologies beyond counterfactual analysis is crucial for achieving a more comprehensive and nuanced understanding of program outcomes. While counterfactual analysis is invaluable for measuring the direct impact of an intervention by comparing it to a scenario in which the intervention did not occur, it has inherent limitations. For instance, it may overlook broader systemic effects or fail to capture the full spectrum of outcomes in complex, multi-faceted programmes. Also, it is not a feasible and suitable methodology for the evaluation of all types of interventions.

By advocating for the use of additional methodologies, such as qualitative research, participatory evaluations, case studies, theory-based evaluations and mixed-method approaches, the guidelines can enable a more holistic assessment of interventions. These alternative methods are particularly effective in evaluating outcomes that are difficult to quantify through counterfactual analysis alone, such as social inclusion, behavioural change, and long-term sustainability. Moreover, they provide richer contextual insights that are essential for understanding the underlying reasons for a program's success or failure, thereby offering valuable lessons for improving future programme design and implementation.

Promoting a wider array of evaluation techniques also enhances flexibility, allowing evaluators to select the most suitable methods based on the specific characteristics of the programme and the context in which it operates (see also paragraph 2.5). This diversified approach ensures that evaluations are more relevant and useful, leading to findings that are not only more insightful but also actionable. Ultimately, this will contribute to more informed, evidence-based policymaking at both national and EU levels.

3.8. Communication and dissemination of evaluation results

To fully leverage the insights gained from evaluations, it is essential that guidelines for evaluations include robust strategies for communicating findings and popularizing results. These recommendations can be provided through a menu of recommended strategies that could be chosen depending on the evaluation's context. This approach enhances transparency, promotes stakeholder engagement, supports evidence-based decision-making, and facilitates the widespread sharing of best practices. It is necessary for the Commission to ensure better cooperation between the units responsible for evaluation and the units responsible for visibility and communication, both at the EU-level as well as member state level.

Some MAs/MS already incorporate robust communication strategies to effectively disseminate evaluation results, ensuring that insights gained have a real impact on policy and practice. It is important that the guidelines for the post-2027 period describe how such strategies can be implemented, allowing all Member States to benefit from the approach.

By sharing evaluation findings in ways that are accessible and relevant to various stakeholders, countries can enhance transparency, encourage broader engagement, and support evidence-based decision-making.

Effective communication strategies can significantly increase stakeholder engagement. Presenting evaluation results through diverse and interactive formats – such as detailed reports, executive summaries, infographics, webinars, and interactive dashboards – can cater to different audience needs and preferences. In this way, Member States can maximize the value of evaluations, turning lessons learned into concrete actions that improve future policy design and implementation.

Popularizing evaluation results can be achieved through various channels, including media releases, policy briefs, and public presentations. By raising awareness about the outcomes and impacts of evaluations, programmes can generate broader interest and engagement. Popularizing results ensures that evaluations reach a wider audience, which can drive greater public and political support for evidence-based initiatives. This approach helps to underscore the importance of evaluations in shaping effective policies and can mobilize additional resources and commitment to implementing recommended changes.

3.9. Incorporating a management response mechanism to evaluation findings

For ensuring that lessons learned from past initiatives are integrated into future programme design and implementation and for improving the accountability and effectiveness of EU-funded programmes it is crucial that after evaluations are completed, findings are not only acknowledged by the MA/MS but, where necessary, also lead to concrete actions that address identified issues and enhance programme performance. Guidelines should recommend a non-binding management response, timeline, and a follow-up plan to address the evaluation findings.

Guidelines could recommend an agreed timeframe for preparing management responses (for all recommendations – accepted and not accepted). Timely preparation of follow-up plans would allow MAs/MS to quickly translate evaluation findings into actionable steps that improve the programme's performance. This responsiveness also would ensure that evaluations remain a relevant tool for continuous improvement rather than becoming a formal exercise with little practical impact.

Member States may choose to create a transparent system for tracking the progress of implementing evaluation findings. This system would provide visibility up to a scope chosen by the Member States, making the management response visible not only to MAs/MS but also to other stakeholders, including the European Commission, civil society, and beneficiaries. Such a system would enhance trust and accountability by offering clear insights into how recommendations are being addressed and into the progress of follow-up actions.

High quality evaluations of key priorities and a formalized, transparent, and time-bound management response mechanism would ensure that evaluation results lead to meaningful improvements. It would also foster a culture of continuous learning and improvement, encouraging more effective management of EU funds. By planning these management response procedures in the evaluation process and indicating them in the evaluation plan it would become a stronger driver for change, ensuring that the recommendations are fully implemented and that their benefits are felt across

programmes. The institutionalized management response is not to be considered as mandatory for the MAs/MS.

3.10. Maintaining independence of evaluation units and experts

Establishing guidelines for qualified personnel and their required skills ensures that MAs/MS are equipped to conduct and/or commission comprehensive and high-quality evaluations (please also see paragraph 2.4). Properly staffed teams with the right research expertise help to avoid delays and deficiencies, resulting in timely, actionable findings that can significantly influence decision-making and enhance program effectiveness.

Equally important is the independence of evaluation experts, which is fundamental to maintaining the credibility and objectivity of the evaluation process. Independent experts who are not involved in the programming or implementation of the programme provide unbiased and impartial assessments. This separation helps to avoid conflicts of interest and ensures that evaluations reflect an accurate and trustworthy analysis of programme performance and impact.

3.11. Differentiate evaluation guidelines from the Better Regulation guidelines

To avoid confusion, it is essential to clearly explain the difference with Better Regulation Guidelines, which is more focused on regulatory impact assessment methodologies rather than evaluation. Regulatory impact assessments serve to anticipate and mitigate potential impacts before new policies are enacted. Some policy makers, who tend to be more used to the impact assessment, might consider it to be sufficient. By clarifying the distinction, the evaluation guidelines will help policymakers to recognize the complementary yet distinct functions and aims of evaluations and regulatory impact assessments. This will ensure that evaluations are correctly understood as tools for assessing the effectiveness of existing policies and programs and for drawing lessons for improvement of programs.

4. Support from the European Commission

4.1. Commission's expectations for evaluations should be made more explicit

Clarifying the European Commission's expectations for Member States' evaluations, as a main stakeholder, is essential for several reasons. Firstly, explicit guidance from the Commission ensures that Member States understand the precise requirements and objectives for evaluations, which helps them in aligning their evaluations with EU-level priorities and standards. Secondly, this ensures that the Commission could more effectively and efficiently benefit from the Member States' evaluations. However, it is important that the aim of clarification does not lead to inflexible and overly narrow regulation, and that the Member States maintain enough liberty to design and conduct evaluations based on their own specific needs.

Additionally, specifying the Commission's expectations helps in the efficient synthesis of national evaluations using meta-analyses, rather than the Commission doing a new evaluation at EU-level. When the Commission's requirements are well-defined, it becomes easier to aggregate findings from different Member States into a cohesive and comprehensive overview. This aggregation avoids duplication of efforts, it maximizes the use of available data, and provides the Commission with a clearer picture of the overall effectiveness of interventions across the EU.

Please also see paragraphs 2.3 and 3.6.

4.2. Revise the approach of conducting synthesis studies

Evaluations conducted by the Commission should focus on specific themes or target group when doing synthesis studies of evaluations carried out by the MAs/MS. This targeted approach allows for a deeper analysis of particular areas of interest or concern, offering more precise insights into the effectiveness of interventions related to those objectives. Additionally, incorporating smaller-scale evaluations conducted by the Commission, such as evaluations covering neighbouring countries, would enhance the learning process and maximize the application of their findings among the Member States.

4.3. Develop evaluation guidelines for the post-2027 period

While the CPR provides the necessary legal foundation, guidelines are crucial for translating these provisions into practical, actionable measures. They ensure effective evaluation, promote consistency across Member States, support ongoing learning and capacity-building, and enhance the trust in the evaluation process. Without guidelines, there is a risk of inconsistencies in the way Member States conduct their evaluations, making it more difficult to compare and synthesize results at the European level.

The guidelines for the post-2027 period should explicitly state they are not mandatory and that the aim of the guidance is to support the evaluation units of the MAs/MS. The guidelines need not be limited to traditional documents; they could take the form of more interactive tools, such as a web portal, providing dynamic support to MAs/MS, as well as could refer to already existing documents.

4.4. Assistance of the Evaluation Helpdesk

We have suggested to apply the OECD evaluation criteria. Nevertheless, we would appreciate the active participation of the Helpdesk. They could assist in clarifying evaluation criteria, as it plays a key role in supporting all MAs/MS. With its broad oversight of both best practices and common pitfalls observed across the different evaluations, they are in an ideal position to provide valuable guidance. This clarification will help ensure that MAs/MS avoid mistakes when establishing their own evaluation standards, ultimately leading to more consistent and effective evaluations.

In general, it is recommended that an evaluation should be conducted after an evaluability assessment. We would appreciate the assistance of the Helpdesk with explaining what it should contain and how it should be carried out. It could be incorporated in the evaluation guidelines.

4.5. Translation of key evaluation reports

To enhance cross-learning among Member States, DG EMPL could translate key evaluation reports from each Member State to English. While summaries in English are frequently provided, making full reports accessible would enable more comprehensive learning and facilitate the sharing of best practices across the EU. The Helpdesk could curate these reports based on criteria such as effective methodology, innovative interventions, and exemplary practices in dissemination. This approach would not only enrich the collective knowledge base but also support the implementation of best practices throughout the EU. The translations should not be a task of the MAs/MS.