1. INTRODUCTION

All evaluations and fitness checks should assess the evaluation criteria of effectiveness, efficiency, coherence, relevance and EU added value of the intervention, or provide due justification why this is not the case. Additional criteria beyond these five can be added.

Evaluations and fitness checks should also always assess the economic, social and environmental impacts of EU interventions (expected or unexpected) with, where relevant, particular emphasis on those impacts identified in a previous impact assessment.

Projects which do not cover all the five evaluation criteria may usefully contribute to a later evaluation (possibly with some updating or confirmation as part of the final process) but are not necessarily considered as "evaluations" unless a prior exception has been authorised.

The degree of analysis conducted for each criterion will depend on the intervention being evaluated, the timing of the evaluation and the reliability of the data (proportionality). Often this will mean that for some criteria new data will need to be collected, analysed and compared with other findings; whilst for others, a short summary can be presented based on existing reports and information.

For example, at an "early" stage in the intervention's lifecycle, it may not be necessary to judge the relevance criterion in any depth. If it is only a few years since the intervention has been made, it may be fair to assume the continued relevance of the action and hence simply restating previous arguments should suffice; alternatively, stakeholder feedback may be the only indicator of whether needs have changed and some summary presentation of their (unchanged) opinions may also be sufficient. Equally, EU-added value may be difficult to judge in the early years, particularly if the early changes are related to setting up new organisations or putting in place a framework; in this instance confirming the validity of the (theoretical) EU added value may be as much as is reasonably possible at that time.

Establishing the intervention logic can be helpful in identifying specific evaluation questions. As mentioned in the tool on designing the evaluation, there is a need to balance generic and specific evaluation questions and to ensure that all questions focus on providing useful information such as information on the changes the intervention sought to achieve, investigating particular intervention characteristics or factors which have / have not worked. It is also advisable not to have too long a list of evaluation questions at the start of an evaluation as it may be too constricting and prevent the analysis from “going where the data leads”. Whilst evaluation sub-questions can be developed early in

---

521 The evaluation of a single intervention may on an exceptional basis omit one or two of the five evaluation criteria. An exception must be granted, and clear justification for such omission must then be provided in the evaluation roadmap and repeated in the final evaluation report. Fitness checks always consider the five criteria.

522 See Tool #1 on Principles, procedures and exceptions.
the evaluation process e.g. to help define a particular question, or drill down on specific areas, this may also happen at a later stage in response to evidence collected.

Depending on the specific evaluation, there may also be overlaps between the criteria and questions asked – for example, in terms of relevance versus EU added value or effectiveness versus impact (sometimes defined as a separate criterion, but given the use within IA of a hierarchy of objectives, considered in the EU context as part of effectiveness). It is generally more important to identify a good set of evaluation questions, than to be too concerned about how to categorise the question, so long as the different criteria are addressed.

It is important to discuss and fix the evaluation questions with the interservice group. Since the evaluation questions do not need to be set at the time that the roadmap is published, they should be discussed at the first meeting and ideally fixed shortly thereafter. These questions will then influence the further design and conduct of the subsequent steps, and be answered in the staff working document.

Figure 1: Simplified view of the intervention and the 5 key evaluation criteria

2. **Effectiveness**

Effectiveness analysis considers how successful EU action has been in achieving or progressing towards its objectives. The evaluation should form an opinion on the progress made to date and the role of the EU action in delivering the observed changes. If the objectives have not been achieved, or things are not on track, an assessment should be made of the *extent* to which progress has fallen short of the target and what factors have influenced *why* something hasn’t been successful or *why* it has not yet been achieved. To this end, effectiveness analysis should seek to identify the factors driving or hindering progress and how they are linked (or not) to the EU intervention.
Consideration should also be given to whether the objectives can still be achieved on time or with what delay. The analysis should also try to identify if any unexpected or unintended effects have occurred. Again, it is important to keep in mind that the evaluation is assessing the role played by the EU intervention – so analysis also needs to consider how the observed changes may be linked to the actions triggered by the EU intervention.

**Box 1: Typical examples of effectiveness questions**

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

In many cases, performance can be identified from monitoring data covering the relevant period. This can then be compared to relevant points of comparison such as:

- The (evaluation) baseline.
- Expectations based on what was expected to have happened at this point (the adopted proposal, drawing on the appropriate information from a prior impact assessment or other relevant documents). This may include comparison with the stated specific / operational objectives.
- Another scenario or benchmark.

3. **Efficiency**

Efficiency considers the relationship between the resources used by an intervention and the changes generated by the intervention (which may be positive or negative). Differences in the way an intervention is approached and conducted can have a significant influence on the effects, making it interesting to consider whether other choices (e.g. as demonstrated via different Member States) achieved the same benefits at less cost (or greater benefits at the same cost).

It is important to note that efficiency analysis should always look closely at both the costs and benefits of the EU intervention as they accrue to different stakeholders. As a general rule, the benefits of EU interventions are expected to justify the costs they generate, although those who bear the costs do not always reap the benefits. This is often the case for safety, health, environment or consumer protection policies. Again, it is

523 See Tool #46 on Designing the evaluation.
524 See also Chapter 8 Methods, models and costs and benefits
important to identify what factors are driving these costs/benefits and how these factors relate to the EU intervention.

Typical efficiency analysis will include analysis of administrative and regulatory burden and look at aspects of simplification – which is important for ALL evaluations. Where appropriate, evaluation findings should pin-point areas where there is potential to reduce inefficiencies, particularly unnecessary regulatory costs, and simplify the intervention. The full efforts to support and perform an intervention can be broken into different categories such as: staff, purchases made, time and/or money spent, fixed costs, running costs, etc. These costs can be associated to different aspects of an intervention and judged against the benefits achieved.

Better regulation and particularly the REFIT programme place a strong emphasis on identifying and where possible measuring (i.e. if possible, quantifying and/or monetising) the costs and benefits of EU interventions.

Sectoral fitness checks pose a specific challenge, because costs corresponding to the examined interventions normally remain with the sector (if not passed on downstream in the form of higher prices), whereas diffuse benefits are enjoyed by society at large. When evaluating the costs/benefits to the sector, it may not be enough to just look at the net position of the sector and see if that is in line with expectations. Although it is beyond the scope of the sectoral fitness check to perform a full evaluation of those acts, some qualitative data on their wider performance have to be considered in order to provide the appropriate context. This is important, because to a certain extent it was this wider context that was used to justify the EU intervention.

A cumulative cost assessment (CCA), although providing important inputs into the evaluation process, is not sufficient on its own to provide the required full picture in terms of the efficiency of the EU intervention. As such, it cannot be "the sole basis for policy recommendations." To serve as an instrument for the policymakers, CCAs need to be put in the context and CCAs need to be supplemented by the analysis of the corresponding benefits arising from the EU legislation.

---

525 See Tool #58 on Typology of costs and benefits

526 See Tool #2 on The Regulatory Fitness programme and the REFIT platform; and Tool #59 on Methods to assess costs and benefits.

527 CCA are studies that aim to estimate the overall regulatory burden on a particular sector. In the EU context, the CCAs will look at the costs arising from the EU regulations.

528 Page 15, Regulatory Fitness and Performance: State of Play and Outlook COM (2014) 368 final

529 Further discussion on costs and benefits analysis in the context of CCA is provided in the Tool #59 on Methods to assess costs and benefits.
Table 1: Approach towards efficiency analysis in specific cases

<table>
<thead>
<tr>
<th>Costs</th>
<th>Benefits</th>
<th>Benefits and costs have same scope and are directly comparable?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specific to the sector</td>
<td>To all affected sectors</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Fitness check</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sectoral fitness check</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Cumulative cost assessment</td>
<td>Yes</td>
<td>Possibly</td>
</tr>
</tbody>
</table>

Assessing costs and benefits may be (methodologically) easier for spending programmes which have well defined stakeholders, systems etc. Doing this with precision at EU level can be difficult since obtaining robust, good quality data to use in the evaluation of costs and benefits may be challenging, particularly across all Member States which may have implemented legislation in a variety of different ways and at different points in time. However, sufficient efforts should be dedicated to this task, given its importance.

**Box 2. Typical examples of efficiency questions**

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

Points of comparison to consider include:

- The (evaluation) baseline.

---

530 See Tool #46 on Designing the evaluation
• What was expected to have happened at this point (the adopted proposal, drawing on the appropriate information from a prior impact assessment or other relevant documents).

• The previous system (e.g. an earlier programme).

• Other similar situations which can be used as benchmarks (e.g. standard IT costs; training or performance benchmarks from private sector, performance management systems in other countries, growth rates, education rates, CO2 levels etc.).

4. **Relevance**

Relevance looks at the relationship between the needs and problems in society and the objectives of the intervention and hence touches on aspects of design. Relevance analysis also requires a consideration of how the objectives of an EU intervention (legislative or spending measure) correspond to wider EU policy goals and priorities. Analysis should identify if there is any mismatch between the objectives of the intervention and the (current) needs or problems. For example, the wrong "problem drivers" may have been identified during the impact assessment; incorrect assumptions may have been made about the cause and effect relationships; circumstances may have changed and the needs/problems now are not the same as the ones looked at when the intervention was designed.

Relevance analysis is very important – because if an intervention does not help to address present needs or problems then it does not matter how effective, efficient or coherent it is – it is no longer appropriate (this is why relevance is sometimes called the "kill" criterion!). This is key information that will assist policy makers in deciding whether to continue, change or stop an intervention and also explains the strong link between relevance and the criterion of EU added value – which assesses whether action continues to be justified at the EU level.

In areas of EU exclusive competence (e.g. trade agreements), the analysis of the relevance and efficiency of the intervention may form a large part of the analysis of EU added value (see EU added value below).

**Box 3. Typical examples of relevance questions**

- To what extent is the intervention still relevant?
- To what extent have the (original) objectives proven to have been appropriate for the intervention in question?
- How well do the (original) objectives of the intervention (still) correspond to the needs within the EU?
- How well adapted is the intervention to subsequent technological or scientific advances? (N.B. Could include issues related to the specify policy here e.g. social, environmental or to implementation, reporting and compliance)
- How relevant is the EU intervention to EU citizens?
Points of comparison\textsuperscript{531} for relevance are usually more qualitative. In terms of the evaluation baseline, it is necessary to think about what were the needs and objectives behind the EU intervention and compare them to the current situation. There may also be grounds to consider how something is expected to be in the near future (based on anticipated or upcoming technological/social/economic changes). For example, in the telecommunications market, discussions about 5G standardisation are well advanced, even if such technology is not in common use at this point.

5. **COHERENCE**

The evaluation of coherence involves looking at how well or not different actions work together. It may highlight areas where there are synergies which improve overall performance or which were perhaps not possible if introduced at national level; or it may point to tensions e.g. objectives which are potentially contradictory, or approaches which are causing inefficiencies.

Checking "internal" coherence means looking at how the various components of the same EU intervention operate together to achieve its objectives e.g. the different articles of a piece of legislation, different actions under an action plan. Similar checks can be conducted in relation to other ("external") interventions, at different levels: for example, between interventions within the same policy field (e.g. a specific intervention on drinking water and wider EU water policy) or in areas which may have to work together (e.g. water policy and chemicals policy, or chemicals and health and safety). At its widest, external coherence can look at compliance with national policies or international agreements/declarations (for example EU labour market interventions might be looking into coherence with ILO conventions) or EU interventions in developing countries.

The focus on coherence may vary depending on the type of evaluation and is particularly important in fitness checks, where coherence analysis will look for evidence of synergies or inconsistencies between actions in a related field which are expected to work together. Even when evaluating an individual intervention, it can be important to check coherence with (a limited number of) other interventions.

When assessing coherence, comparison with other scenarios is again likely to be predominantly qualitative. Suitable points of comparison\textsuperscript{532} could include looking at changes in coherence between the start and end of the period being evaluated or information from the impact assessment. They might also include looking at existing or new national or international actions. The level of coherence being evaluated (e.g. internal to a given EU action, within a given policy field, wider EU policy or wider global context) could also affect the comparison point and degree of analysis possible.

**Box 4. Typical examples of coherence questions**

- To what extent is this intervention coherent with other interventions which have similar objectives?

\textsuperscript{531} See Tool #46 on Designing the evaluation.

\textsuperscript{532} See Tool #46 on Designing the evaluation.
To what extent is the intervention coherent internally?
To what extent is the intervention coherent with wider EU policy?
To what extent is the intervention coherent with international obligations?

6. EU-ADDED VALUE

EU-added value\textsuperscript{533} looks for changes which it can reasonably be argued are due to the EU intervention, over and above what could reasonably have been expected from national actions by the Member States. In many ways, the evaluation of EU added value brings together the findings of the other criteria, presenting the arguments on causality and drawing conclusions, based on the evidence to hand, about the performance of the EU intervention.

Under the principle of subsidiarity (Article 5 Treaty on European Union), and in areas of non-exclusive competence, the EU should only act when the objectives can be better achieved by Union action rather than action by the Member States. It requires consideration of the added value of EU action compared to that of other actors (see also Tool #5 on Legal basis, subsidiarity and proportionality). EU added value analysis should, where applicable, respond to the subsidiarity analysis conducted in any related IA.

The sources and nature of this additional value vary from intervention to intervention. It is, in particular, useful to distinguish the European added value of an EU policy measure in general (such as an EU regulation to foster the single market) and that of an EU spending programme per se. In both cases, European added value may be the results of different factors: coordination gains, legal certainty, greater effectiveness or efficiency, complementarities etc. In all cases, concluding on the continued need for the intervention at EU level may be difficult as the measurement of EU added value is challenging.

In areas where the EU has exclusive competence, the appropriate answer to the question of EU added value may simply involve re-stating the reasons why the EU has exclusive competence or may already be answered by the efficiency and effectiveness analysis.\textsuperscript{534}

In the context of the EU budget, the Commission staff working documents SEC(2011) 867 final and SWD(2015) 124 final recommend that the EU added value test is performed on the basis of the following 3 criteria:

**Effectiveness:** where EU action is the only way to get results to create missing links, avoid fragmentation, and realise the potential of a border-free Europe.

**Efficiency:** where the EU offers better value for money, because externalities can be addressed, resources or expertise can be pooled, an action can be better coordinated.

**Synergy:** where EU action is necessary to complement, stimulate, and leverage action to reduce disparities, raise standards, and create synergies.

---

\textsuperscript{533} For further information see SEC(2011) 867 final "The added value of the EU budget".

\textsuperscript{534} See Tool #5 on Legal basis, subsidiarity and proportionality.
As (big) data becomes easier to collect, possibilities grow to apply new methods to analyse the EU added value (e.g. counterfactual impact evaluation). Where there are difficulties identifying a robust counterfactual, the analysis of EU added value should as a minimum provide qualitative, reasoned arguments about the likely role/contribution of the EU intervention, backed by appropriate quantitative and qualitative evidence. It is also important that evaluations clearly state the challenges that have been encountered and resulting limitations in the certainty or accuracy of such findings, which can vary greatly from case to case.

**Box 5. Typical questions on EU added value**

- What is the additional value resulting from the EU intervention(s), compared to what could reasonably have been expected from Member States acting at national and/or regional levels?  

- What would be the most likely consequences of stopping or withdrawing the existing EU intervention?

When assessing the EU added value, the comparison is likely to involve consideration of performance against both the (evaluation) baseline and, if available a projection of how the situation was expected to evolve without the EU intervention (a defined counterfactual, or some estimate of the cost of the Union not acting - "the cost of non-Europe"). Often such analysis is qualitative, analysing whether the subsidiarity arguments put forward before the intervention (as presented in a prior impact assessment, or other accompanying documents) were valid and whether the expected changes resulting from EU action were delivered. It may also be appropriate to analyse whether any contextual change, or other factors affected the assumption that such change could only be generated by EU level action.

7. **OTHER EVALUATION CRITERIA**

There are also several further evaluation criteria which it may be appropriate to consider, depending on the type of intervention and the timing of the evaluation. The most common additional criteria evaluated by the Commission are shown below.

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

**Complementarity:** To what extent do EU policies and interventions support and usefully supplement other policies (in particular those pursued by the Member States)?

---

535 In answering this question, it may be interesting to consider the question for each criteria. For example, How efficient has the EU intervention been in comparison to what could reasonably have been expected from Member States acting at national and/or regional levels?

536 See Tool #46 on Designing the evaluation.
Coordination: To what extent are interventions organised to maximise their joint effects, e.g. by mobilising resources combined with harmonising measures?

Equity: how fairly are the different effects distributed across the different stakeholders / regions? / genders? / Social groups?

Sustainability: How likely are the effects to last after the intervention ends? It is often hoped that the changes caused by an intervention are permanent. It can be important to test this expectation for interventions which have a finite duration, such as particular programmes.

Acceptability: To what extent can we observe changes in the perception of the intervention (positive or negative) by the targeted stakeholders and/or by the general public?

8. Good Practice Tips

- Evaluation questions should be worded in a way that forces the evaluator to go beyond an answer based on simple description. Questions that start with How, Why, To what extent are more likely to ensure that the answer provided looks at what the links were between the changes observed and the EU intervention(s). Questions that start with verbs such as "Do" the directives…? "Are" the directives providing…? "Should" be avoided as they tend to provoke yes/no answers.

- Try not to have too many evaluation questions. Sometimes it is necessary to have very specific questions, other times it is better to have a more generic set and see where the data/analysis leads. There is always a trade-off between the number of questions that are set and the depth of analysis that can be conducted, especially across all Member States.

- For spending programmes, it may be necessary to assess to what extent has it been possible to prevent and detect fraud.

- Check any prior impact assessment to see what issues were addressed and what expectations were presented. Where necessary compare the proposal accompanying the impact assessment to the final actions adopted/introduced and try to identify where amendments to the Commission's proposal may have changed the intervention logic described in the impact assessment.

- Encourage consideration of the "end-user" perspective. End-users are most affected by actions triggered due to EU interventions – they have practical experience of what has happened on the ground and may have a different perspective from policy makers, governments, NGOs etc.

- Evaluation questions are often worded in a technical way, using terminology that makes sense within the Commission but which requires further explanation (simplification) to many stakeholders. "Translate" the criteria into more manageable / understandable concepts, particularly for consultation and data gathering exercises. For example:
  - Rather than referring in an abstract manner to the "objectives of the policy" describe the objective in more common terms – so ask "What progress has been
made towards increasing the availability of funding to small and medium sized businesses?"

• If there is a problem with the timetable and it becomes clear that any deadlines set for the final evaluation will be missed, it makes sense to let interested parties know. This is particularly true where an evaluation is set in a legal act and the evaluation findings should be communicated to Parliament and Council.

• The exact scope has significant impacts on the final design. If there are issues about what EU actions to include in scope (this is of particular relevance for fitness checks), consider doing a short "scoping" exercise first.