

# Executive summary

The Organisation for Economic Co-operation and Development (OECD) has established common definitions for six evaluation criteria – relevance, coherence, effectiveness, efficiency, impact and sustainability – to support consistent, high-quality evaluation. These criteria provide a normative framework used to determine the merit or worth of an intervention (policy, strategy, programme, project or activity). They serve as the basis upon which evaluative judgements are made.

The criteria can be thought of as a set of lenses, providing complementary perspectives that together give a holistic picture of an intervention and its results. The criteria encourage evaluators – as well as those involved in designing or managing interventions – to think more deeply about the nature of an intervention, its implementation process and its results.

Two guiding principles were set out by the OECD DAC Network on Development Evaluation alongside the definitions of the six criteria. These are:

1. Principle One: The criteria should be applied thoughtfully to support high-quality, useful evaluation.
2. Principle Two: Use of the criteria depends on the purpose of the evaluation.

The criteria should be applied thoughtfully and adapted to the context of the intervention and the evaluation. Though originally developed for use in international development co-operation, the criteria can be applied in any sector and for evaluations of public or private interventions. They can be used to evaluate many different topics and types of interventions – including thematic topics or strategic issues, policies and projects. Different methodologies can be used when evaluating based on the criteria, recognising that the criteria form just one part of the full spectrum of evaluation norms and standards.

Evaluators should work in ways that thoughtfully consider differential experiences, and reflect how power dynamics based on gender and other forms of discrimination (e.g. age, race/ethnicity, social status, ability) affect the intervention's implementation and results.

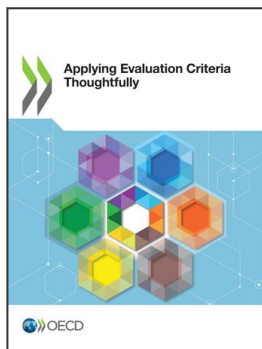
The key concepts for each of the criterion are defined as follows:

- Relevance entails examining the extent to which the intervention's objectives and design respond to beneficiaries' needs and priorities, as well as alignment with national, global and partner/institutional policies and priorities. Understanding gendered power dynamics and reflecting on the commitment to "leave no one behind" are crucial in understanding relevance. If circumstances change, evaluations should also look at whether interventions remain relevant.
- Coherence – the new criteria – examines the extent to which other interventions (particularly policies) support or undermine the intervention and vice versa. This includes internal coherence (within one institution or government) including compatibility with international norms and standards, and external coherence (with other actors' interventions in the same context). Coherence includes concepts of complementarity, harmonisation and co-ordination, and the extent to which the intervention is adding value while avoiding duplication of effort. In line with the 2030 Agenda, greater attention must be paid to coherence, with an increased focus on the

synergies (or trade-offs) between policy areas. This new criterion encourages an integrated approach to understanding complex interventions and their results.

- Effectiveness looks at the extent to which the intervention achieved, or is expected to achieve, its objectives and its results, while taking into account the relative importance of the objectives. The new definition encourages analysis of differential results across groups and the extent to which the intervention contributes to or exacerbates equity gaps. Effectiveness is the most commonly evaluated criteria and is often used as an overall measure of success.
- Efficiency helps evaluators ask questions about the extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way. “Economic” is the conversion of inputs (funds, expertise, natural resources, time, etc.) into results, in the most cost-effective way possible, as compared to feasible alternatives in the context. The new definition includes the dimension of “timely delivery”. This criterion is an opportunity to check whether an intervention’s resources can be justified by its results, which is of major practical and political importance. Many stakeholders, including beneficiaries, care about efficiency, because it can support better use of limited resources, to achieve more.
- Impact is the extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects. Impact addresses the intervention’s ultimate significance and potentially transformative effects – holistic and enduring changes in systems or norms. The impact criterion goes beyond effectiveness and encourages consideration of the big “so what?” question. This is where evaluators look at whether or not the intervention created change that really matters to people.
- Sustainability is the extent to which the net benefits of the intervention continue or are likely to continue. Depending on the timing of the evaluation, this may involve analysing the actual flow of net benefits or estimating the likelihood of net benefits continuing over the medium and long term. While the underlying concept of continuing benefits remains unchanged, the new definition encompasses several elements for analysis – financial, economic, social and environmental – and attention should be paid to the interaction between them.

A variety of challenges may be faced when applying each of the criteria, including common issues related to weaknesses in intervention design and data availability. Practical suggestions are provided for how evaluators and evaluation managers can overcome these challenges, including by assessing evaluability early in the process, working with stakeholders to document decision made or recreate baselines, and being transparent about any limitations to manage expectations. These tables – and the rest of the guidance – will be updated as further lessons from experience are gathered.



**From:**  
**Applying Evaluation Criteria Thoughtfully**

**Access the complete publication at:**  
<https://doi.org/10.1787/543e84ed-en>

**Please cite this chapter as:**

OECD (2021), "Executive summary", in *Applying Evaluation Criteria Thoughtfully*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/e3a75dff-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.