



## EVALUATION CRITERIA & QUESTIONS

SIMONE BUSETTI  
[sbusetti@unite.it](mailto:sbusetti@unite.it)

# CONTENTS OF THIS CLASS

- Evaluation criteria
- Formulating evaluation questions

# EVALUATION CRITERIA

# EVALUATION CRITERIA

Evaluation criteria have a normative role: they describe the desired attributes of interventions

They should be contextualised to the intervention, evaluation, stakeholders, goals, etc.

Evaluation on these criteria fulfils both accountability and learning goals

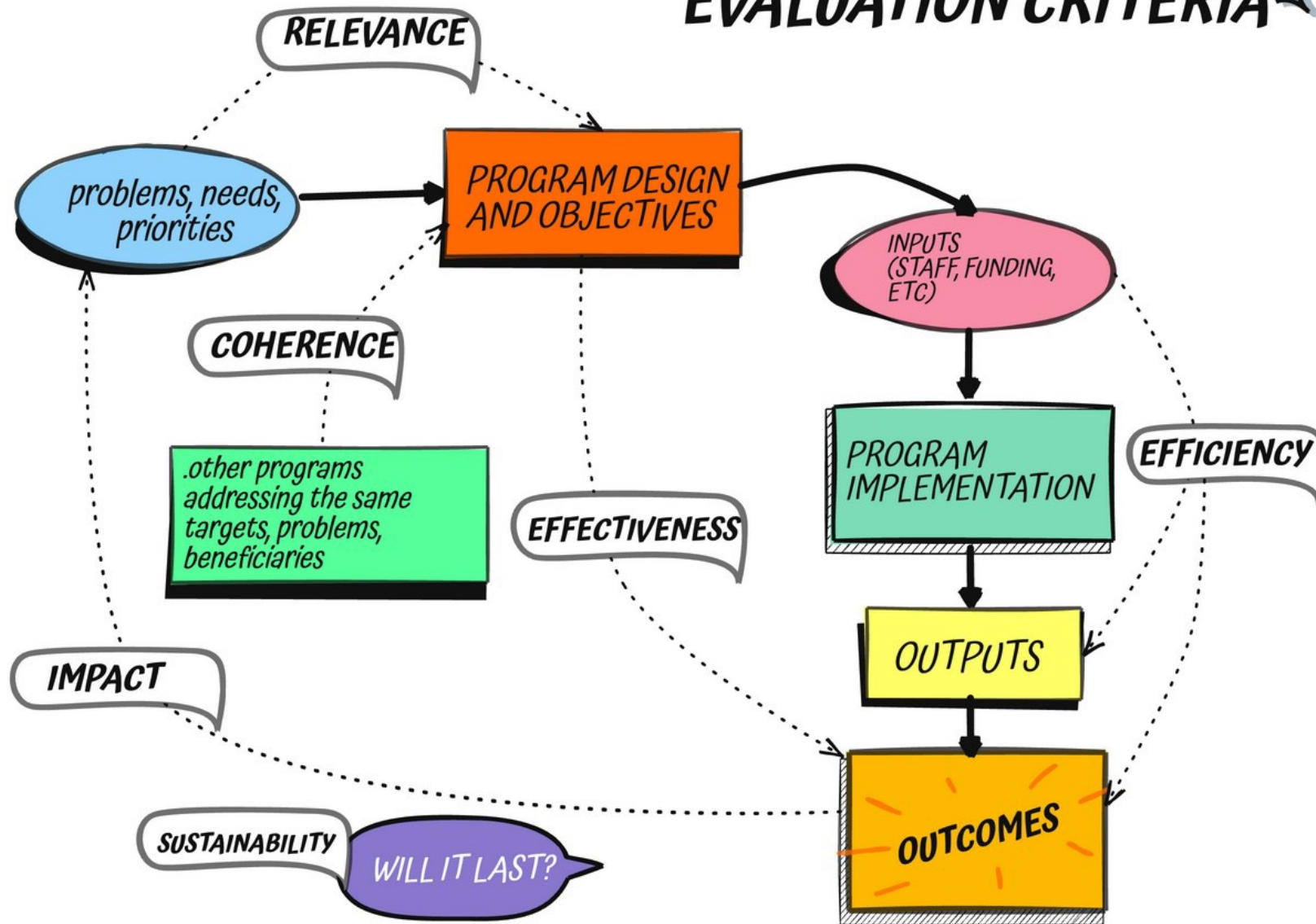
See

[https://www.oecd.org/en/publications/better-criteria-for-better-evaluation\\_15a9c26b-en.html](https://www.oecd.org/en/publications/better-criteria-for-better-evaluation_15a9c26b-en.html)

OECD SIX EVALUATION CRITERIA:

1. RELEVANCE
2. COHERENCE
3. EFFICIENCY
4. EFFECTIVENESS
5. IMPACT
6. SUSTAINABILITY

# EVALUATION CRITERIA



# 1. RELEVANCE:

## IS THE INTERVENTION DOING THE RIGHT THINGS?

*The extent to which the intervention objectives and design respond to beneficiaries', global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change.*

It mainly looks at if the intervention is well-designed to address priorities, problems, needs.

It is often evaluated at the time of the intervention design, but also later in time if circumstances change.

Relevance assessments may involve looking at tradeoffs between different priorities and needs.

## 2. COHERENCE:

### HOW WELL DOES THE INTERVENTION FIT?

*The compatibility of the intervention with other interventions in a country, sector or institution.*

One can evaluate *internal coherence* (with respect to other programs by the same institution) or *external coherence* (with respect to other actors' programs).

Bad coherence is inherent in duplications, overlaps, conflicting interventions, lack of synergies or complementarities, etc.

Focusing on coherence is a way to take a 'system view' of the intervention, i.e. its contribution in a wider set of activities and institutions.

### 3. EFFICIENCY:

## HOW WELL ARE RESOURCES BEING USED?

*The extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.*

Economic means the conversion of inputs (funds, expertise, natural resources, time, etc.) into *outputs, outcomes and impacts*, in the most effective way possible as *compared to feasible alternatives*.

Timely means within the intended time or a reasonable timeframe, depending on context.

Consider that the implementation of a program—how results are produced, i.e. the program operations and management—are fundamental to a measurement of efficiency.



## 4. EFFECTIVENESS:

### IS THE INTERVENTION ACHIEVING ITS OBJECTIVES?

*The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups.*

Effectiveness looks at closely attributable results (proximal outcomes directly derived from the intervention).

You may focus on objectives and/or results and many times you will need to select which to measure.

Pay close attention to heterogeneity of results across places, administrations, population subgroups, etc.

## 5. IMPACT:

### WHAT DIFFERENCE DOES THE INTERVENTION MAKE?

*The extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects.*

Ultimate significance and transformative effects in a social, environmental and economic perspective, longer term and broader scope effects not captured by effectiveness; indirect, secondary and potential consequences.

## 6. SUSTAINABILITY: WILL THE BENEFITS LAST?

*The extent to which the net benefits of the intervention continue, or are likely to continue.*

What is the capacity and likelihood of the system to sustain benefits over time, in the medium and long-term?

Sustainability may cover all range of effects across various dimensions (financial, social, environmental, etc.)

# AN EXAMPLE WITH EVALUATION QUESTIONS: THE CASE OF SOLIDARITY MARKETS

## RELEVANCE

To what extent does the design of the program (e.g. food procurement systems) allow for a satisfactory availability of food? How does it promise to respond to beneficiaries' nutritional needs, preferences, and dignity?

## COHERENCE

To what extent does the intervention complement and coordinate with other municipal or third sector programs?

## EFFECTIVENESS

To what extent does the Solidarity market make people more food secure, less stigmatised and more socially included?

## EFFICIENCY

How cost-effectively are financial, material, and volunteer resources used compared to alternative food assistance mechanisms?

## SUSTAINABILITY

To what extent is the intervention and its operations self-sufficient and sustainable in a durable way?

## IMPACT

What broader changes—intended or unintended—has the Solidarity Market generated in participants' lives and the wider community? How?

# AN EXAMPLE WITH RESULTS:

## AN IMAGINARY RURAL CLEAN WATER PROGRAM

### 1. Relevance

The clean water program aligns with the community's priority to reduce waterborne diseases and supports the government's national sanitation policy. It was designed after participatory consultations with local women, who identified a lack of safe water as their top concern.

### 2. Coherence

The program complements the Ministry of Health's hygiene education campaign and avoids duplication with a separate donor-funded borehole project. However, limited coordination with an NGO sanitation initiative led to some overlap in target villages.

### 3. Effectiveness

The project aimed to provide clean water to 20,000 people. By the end of Year 3, 18,500 people had access, reaching 92% of the target. Reported cases of diarrheal disease declined by 35% in the project area.

### 4. Efficiency

The average cost per functioning water point was USD 8,000—15% below the regional benchmark. However, delays in procurement caused a 6-month extension, increasing administrative costs.

### 5. Impact

Beyond improved health outcomes, women reported saving two hours per day previously spent fetching water, allowing more time for income-generating activities. On the downside, some wells led to overuse of groundwater, slightly lowering the water table in dry months.

### 6. Sustainability

Water user committees were trained and have successfully collected maintenance fees for one year after project closure. However, technical capacity for pump repairs remains weak, posing a risk to long-term sustainability.

# **FORMULATING EVALUATION QUESTIONS**

# GOOD PRACTICE TIPS — BETTER REGULATION TOOLBOX

[HTTPS://COMMISSION.EUROPA.EU/LAW/LAW-MAKING-PROCESS/BETTER-REGULATION/BETTER-REGULATION-GUIDELINES-AND-TOOLBOX EN](https://commission.europa.eu/law/law-making-process/better-regulation/better-regulation-guidelines-and-toolbox_en)

Word the questions to go beyond simple descriptions. 'How', 'Why', 'To what extent' call for nuanced answers about processes, non-program factors, and attribution. 'Did the intervention...' or 'Is the intervention...' call for yes/no answers.

If necessary, use clarifications or sub-questions to define specific elements to consider (e.g. in answering the question, special attention should be paid to...)

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 and the depth of analysis that can be conducted, especially across different cases.

Check any prior evaluation to see what issues were addressed and what expectations were presented. Copying is a good idea...

Encourage consideration of the 'end-user' perspective; they are most affected by the interventions, have practical on-the-ground experience, and may have a different perspective.

Evaluation questions should avoid technical jargon; they should 'translate' the criteria into more manageable and understandable concepts.

Dare to ask the difficult question, including those that may challenge current policy.

Ask for lessons! Where do we go from here?

# EXAMPLES OF EFFECTIVENESS QUESTIONS

How successful has the intervention been in achieving (or progressing towards) its objectives? To what extent were the expected changes resulting from the action delivered?

What have the quantitative and qualitative effects of the intervention been? Are they as expected when introducing the intervention?

What external factors have affected progress towards the objectives and how are they linked to the intervention?

If the objectives have not been achieved yet, can the objectives still be achieved in time? If not, with what delay?

Are there any unexpected or unintended effects that have occurred, and which drove or hindered progress? What can explain these effects?



# EXAMPLES OF EFFICIENCY QUESTIONS

What are the costs and benefits for different stakeholders? Were the costs and benefits of the intervention distributed as expected among them? What could explain the differences (if any)?

To what extent are the administrative and/or adjustment costs of the intervention justified, given the changes/effects it has achieved?

Have the changes/effects of the intervention been achieved at initially expected costs or were these costs different for whatever reason? What could explain the differences (if any)?

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?

Have any inefficiencies been identified? Could the intervention have been done in a more efficient way? What is the simplification and cost reduction potential of the intervention?

If there are significant differences in costs (and benefits) between territories, sectors, etc., what is causing them? Were they expected?

How timely and efficient is the intervention's administrative process (e.g. for reporting and monitoring)?

# EXAMPLES OF RELEVANCE EX-POST QUESTIONS

To what extent did the scope and objectives of the intervention remain relevant over the implementation period?

How did the objectives of the intervention correspond to broader policy goals and priorities? •

How well do the (original) objectives of the intervention still correspond to the identified needs?

To what extent do the needs/problems addressed by the intervention continue to require action?

How well adapted is the intervention to social, environmental changes or developments related to implementation, reporting and compliance?

Have the circumstances changed so much in the meantime that the intervention has to change or adapt over time?

# EXAMPLES OF COHERENCE QUESTIONS

To what extent is this intervention coherent with other EU, national, and local interventions that have similar objectives?

Does the intervention comply with the green oath to 'do no significant harm', and could it help respond better to significant harm?

To what extent are the various elements of intervention coherent with one another?

To what extent have the various elements of intervention generated synergies and/or compensated possible trade-offs among them?

To what extent is the intervention coherent with (current) wider policies and priorities?

To what extent is the intervention coherent with international obligations, including the SDGs?

# EXAMPLES OF **IMPACT** QUESTIONS

What have been the wider effects of the intervention on society, the economy, or the environment — beyond its immediate objectives?

To what extent has the intervention contributed to long-term or systemic changes (e.g. in behaviours, institutions, or policies)?

Are there any positive or negative unintended impacts (e.g. social, environmental, distributional) that can be linked to the intervention?

Have there been any spill-over effects (across sectors, territories, or target groups)?

How have different stakeholders been affected by these broader changes (positively or negatively)?

What mechanisms or causal pathways explain the observed impacts?

To what extent can the observed impacts be attributed to the intervention, as opposed to other external factors?

How do the observed impacts compare with what was expected when the intervention was designed?

Have there been any multiplier effects, innovations, or changes in public attitudes arising from the intervention?

# EXAMPLES OF SUSTAINABILITY QUESTIONS

To what extent are the results and benefits of the intervention likely to continue after external support ends?

Are the institutional, financial, or social arrangements in place sufficient to maintain the achieved outcomes over time?

Have stakeholders (beneficiaries, institutions, partners) developed the ownership and capacity needed to sustain results?

Are there any risks that could threaten the durability of results (e.g. funding cuts, turnover, policy change, environmental shifts)?

To what extent are results embedded in policy frameworks, local governance, or community practices?

Are there mechanisms for learning, adaptation, and updating to keep the intervention relevant and effective over time?

How dependent is the intervention on external actors or temporary funding? What would happen if that support stopped?

What factors could enhance or undermine the sustainability of the intervention's effects?

# EXERCISE: EVALUATION CRITERIA AND QUESTIONS

## GREENSTEPS — A FICTIONAL VOLUNTEERING PROGRAM

### CONTEXT

The city of Valleverde (150,000 residents) faces low youth civic engagement and neglected public green areas. In 2022, the municipality launched “GreenSteps”, a three-year programme to involve young people (16–25 years old) in cleaning and maintaining parks, while promoting environmental awareness and community belonging. The program is funded by the municipality, with 30% participation from a Bank Foundation.

### OBJECTIVES

- Environmental: Improve the cleanliness and biodiversity of urban green spaces.
- Educational: Strengthen environmental knowledge and sustainable behaviour among youth.
- Social: Increase civic responsibility and teamwork.
- Institutional: Build partnerships between schools, NGOs, and the municipality.

# EXERCISE: EVALUATION CRITERIA AND QUESTIONS

## GREENSTEPS — A FICTIONAL VOLUNTEERING PROGRAM

### PROGRAM ACTIVITIES:

- Recruitment and training through schools and youth groups.
- Teams “adopt” local parks, clean and maintain them.
- Youth-led social media and awareness events.
- The municipality collects data (hours volunteered, waste collected, events held) and awards certificates.

### YOUR TASK:

- Apply the evaluation criteria, formulate relevant questions, and explain why asking those questions on this specific program.