



# Evaluation Standards

**Dr HMP Herath**

**Senior lecturer**

**Ph.D. (USJ), RN(SLNC), B.Sc. (Nursing) Hons(UPSL), CTHE (KDU)**

**Department of Nursing & Midwifery**

**FAHS, KDU**



Co-funded by the  
Erasmus+ Programme  
of the European Union

Reference number: 618596-EPP-1-2020-1-SE-EPPKA2-CBHE-JP

This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.





# Learning Objectives

1. Understand the importance of evaluation criteria in assessing the quality and value of development interventions.
2. Identify and describe the five basic standards for evaluating development interventions: effectiveness, impact, relevance, sustainability, and efficiency.
3. Explain the concept of effectiveness and its measurement at the levels of outputs, outcomes, and impact.
4. Analyze the criteria for evaluating humanitarian assistance, including appropriateness, coverage, connectedness, and coherence.
5. Understand the role of efficiency in evaluating development interventions and distinguish between technical efficiency and allocative efficiency.
6. Recognize the importance of fairness and unbiasedness in evaluation and its relevance in accountability and learning evaluations.



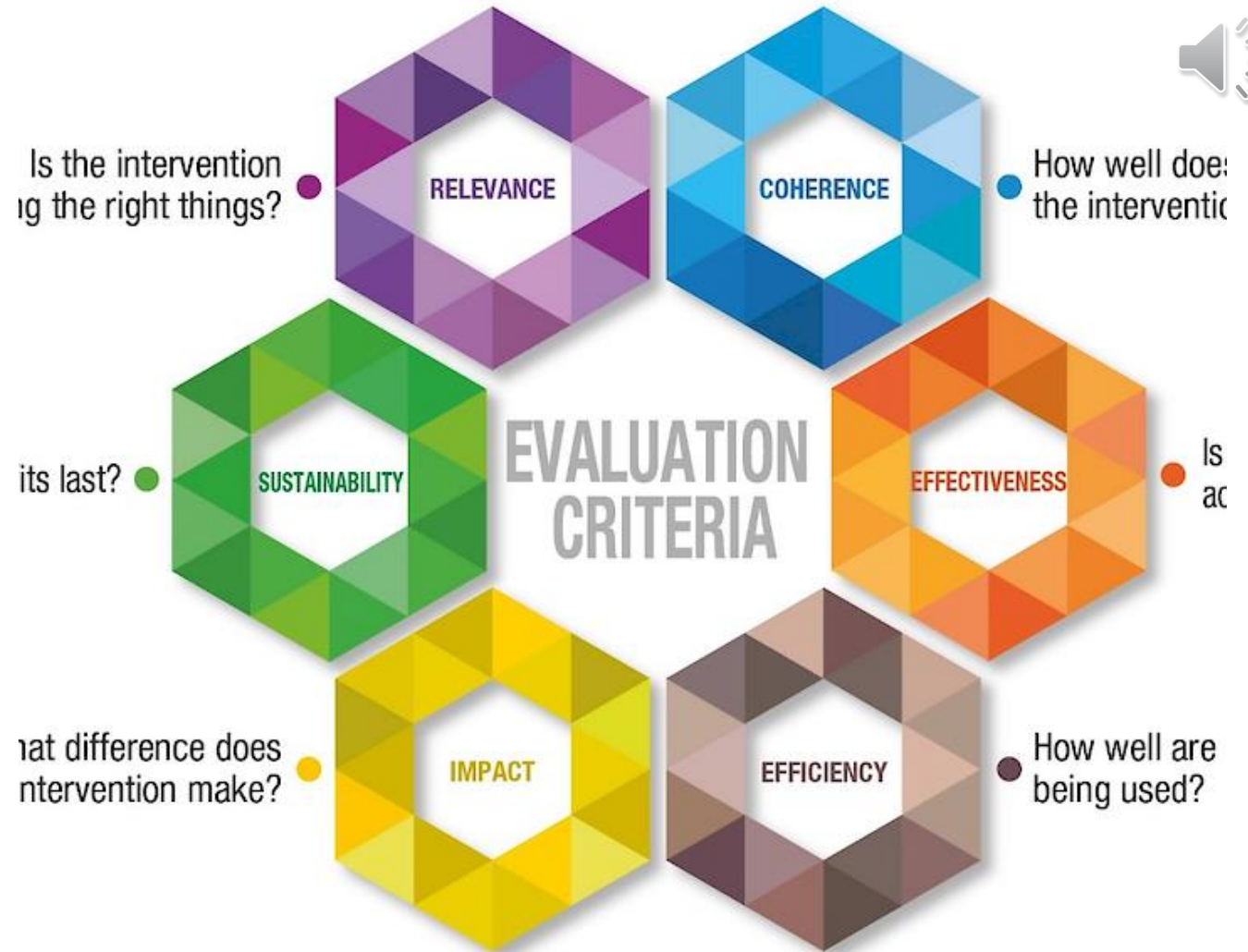
# What is Evaluation?

- “An evaluation is a careful and systematic retrospective assessment of the design, implementation, and results of development activities.”



# Principal evaluation criteria

- Every evaluation has one or more criteria that are used to judge, either explicitly or implicitly, how good or valuable the thing being evaluated is.





# Evaluating development interventions:

Here are the five metrics that are recommended as basic standards for assessing development interventions:

- Effectiveness
- Impact
- Relevance
- Sustainability
- Efficiency



# Evaluating development interventions:



**1. EFFECTIVENESS** - The degree to which a development action has met its goals, taking into account how important each goal was.



**2. IMPACT** - The totality of the effects of a development intervention, positive and negative, intended and unintended



**3. RELEVANCE** - The extent to which a development intervention conforms to the needs and priorities of target groups and the policies of recipient countries and donors.



# Evaluating development interventions:



## 4. SUSTAINABILITY

The continuity or durability of development benefits following the cessation of development assistance.



## 5. EFFICIENCY

Considering alternatives, the extent to which the costs of a development intervention can be justified by its results.





# Effectiveness and Efficiency



Effectiveness refers to the extent to which an evaluated intervention has achieved its objectives.



Efficiency, by contrast, refers to the extent to which the costs of an intervention can be justified by its results.



# Evaluating development interventions

- When we focus on relevance, we assess the evaluated intervention in connection to larger contexts of needs, priorities, and policies, as well as the intervention's aims and methods for attaining them.
- Finally, sustainability is a criterion for assessing the chance that the benefits created by an intervention will be maintained if external support is terminated.



# Further criteria

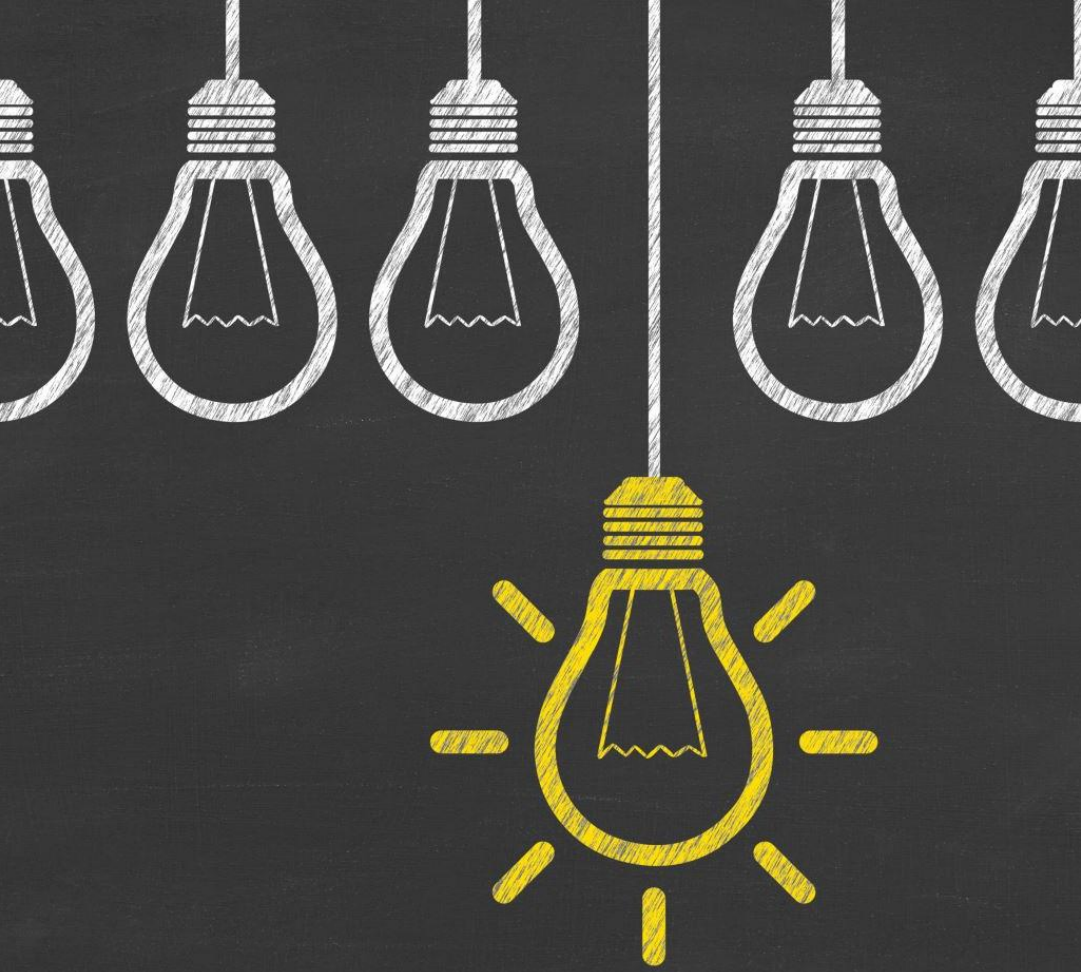
---





## Further criteria

- In many evaluations procedural values and principles are used as evaluation criteria.
- Participation, partnership, human rights, gender equality, and environmental sustainability, are prominent examples.
- They are all values and principles governing the design and implementation of interventions.



## Criteria for evaluation of humanitarian assistance

- Appropriateness
- Coverage
- Connectedness
- Coherence

authors, and the Commission  
the information contained therein.

# Criteria for evaluation of humanitarian assistance



- **APPROPRIATENESS** - The extent of humanitarian inputs and activities are tailored to local needs.
  - The requirements of ownership
  - Accountability
  - cost-effectiveness
  - How do activities respond to the changing demands of the situation?
- **COVERAGE** - The extent to which the entire group in need had access to benefits and was given the necessary support.
  - Did the advantages reach the target population or did too much seep out?
  - Were advantages evenly spread across gender, age, and cultural lines?



# Criteria for evaluation of humanitarian assistance



- **CONNECTEDNESS** - The extent to which short-term emergency activities take into account longer-term needs and the interconnectedness of humanitarian problems.
  - Examples of problems to be dealt with are the environmental effects of refugee camps
  - Damage of roads through food transports
  - Damage to local institutions as a result of international NGOs taking over central government functions
- **COHERENCE** - Consistency between development, security, trade, military and humanitarian policies, and the extent to which human rights were taken into account.
  - Were policies mutually consistent?
  - Did all actors pull in the same direction?
  - Were human rights consistently respected?





# Standards of performance

**It is important to distinguish between evaluation criteria and the performance standards that relate to those criteria.**

- Standards of performance - values on those variables that reflect acceptable levels of achievement
- Evaluation criteria - variables in terms of which performance is measured or judged.

**Example: If we are dealing with a road construction intervention and effectiveness is the criterion of evaluation.**

- The completion of one kilometer of asphalt road every three months, for instance, could serve as a performance criteria



Co-funded by the  
Erasmus+ Programme  
of the European Union

Reference number: 618596-EPP-1-2020-1-SE-EPPKA2-CBHE-JP  
This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.





# Possible issues of performance standards



- Standards too ambitious or not ambitious enough.  
(Evaluators might even propose an alternative set of standards)
- Goals and objectives of development interventions may be vague, incomplete or both.

# Possible issues of performance standards



- However, performance criteria must reflect the goals, commitments, and formal obligations of individuals involved in the reviewed activity and not be arbitrary impositions.
- Insiders may view performance standards set by evaluators or other outsiders as unfair or unimportant.
- Fairness is vital in accountability evaluations but also in learning evaluations.





**Effectiveness**



# Effectiveness



- The term effectiveness refers to the extent to which the objectives of an intervention have been achieved as a result of the implementation of planned activities.
- Effectiveness can be measured at the level of outputs/ at the levels of outcome and impact.
  - First, we are concerned about meeting production targets
  - Second involves achieving the desired consequences of these services and goods.

# Assessing effectiveness



- Most of the time, it's harder to measure success at the level of outcomes than at the level of outputs.
- At the output level, the task is done when the intervention's products and services meet quantity and quality targets.
- In the absence of quality standards, assessing output quality is challenging. However, most problems have solutions.





# Assessing effectiveness

- At the outcome level an assessment of effectiveness is made in two steps.
  - *First, the achievement of objectives is measured.*
  - *Second, the issue of causal attribution is addressed.*
- Both these steps are important.
- In most cases there are many factors at play in addition to the evaluated intervention.
- Therefore, it must be thoroughly considered and ruled out as much as possible that recorded changes would have occurred even without the intervention.



# Assessing effectiveness

- Goal achievement may be tough to measure. Because one or more of the following categories of information may be missing:
  - Background information before the intervention
  - Records of changes compared to baseline during the intervention and afterward
  - An empirical description of the intervention's expected outcome.



# Assessing effectiveness



- Missing baseline data is a common issue.
  - Written documentation and interviews with target groups and others can assist recreate a baseline.
  - Due to human memory, a baseline reconstructed by memory recall is frequently less exact than a baseline assembled before the intervention.
- Vague descriptions of the intervention's intended outcome could cause issues.
  - If goals and objectives are unclear, it may be hard to determine if and how documented changes indicate an achievement.
  - Many initiatives use empirical indicators to specify aims and objectives, although indicators may pose questions. Indicators should never be assumed to be valid and meaningful.
- Objective formulations grow more imprecise as we progress up the hierarchy.





# Assessing effectiveness

- Effectiveness is crucial for assessing development interventions.
- Principals, funders, and others want to know results were achieved as promised for accountability.
- Intervention managers who need the same information for management and learning are equally crucial.
- Setting demanding but achievable goals is a basic management skill.
- If our aims and objectives are never realized, something can be wrong





# Shortcomings of effectiveness

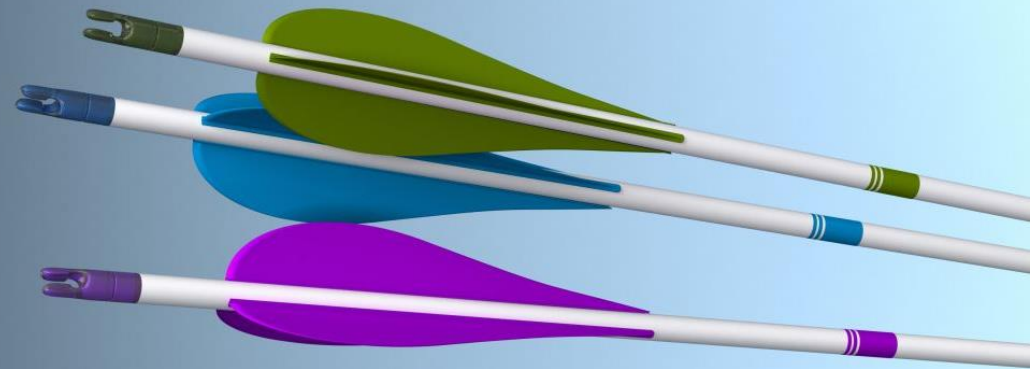
- The assessment of effectiveness has only limited value.
- Can know whether the evaluated intervention has achieved its goals or not.
- Does not address the relevance and usefulness of the results or unexpected impacts.



# Shortcomings of effectiveness

## Standard questions about effectiveness:

- How well do changes in the target area match the expected results, purpose, and goal of the action that was evaluated?
- How much of the recognized change is due to the intervention and how much is due to other things?
- Why are goals met or not met?
- How can the action be made more effective?





# Impact



Co-funded by the  
Erasmus+ Programme  
of the European Union

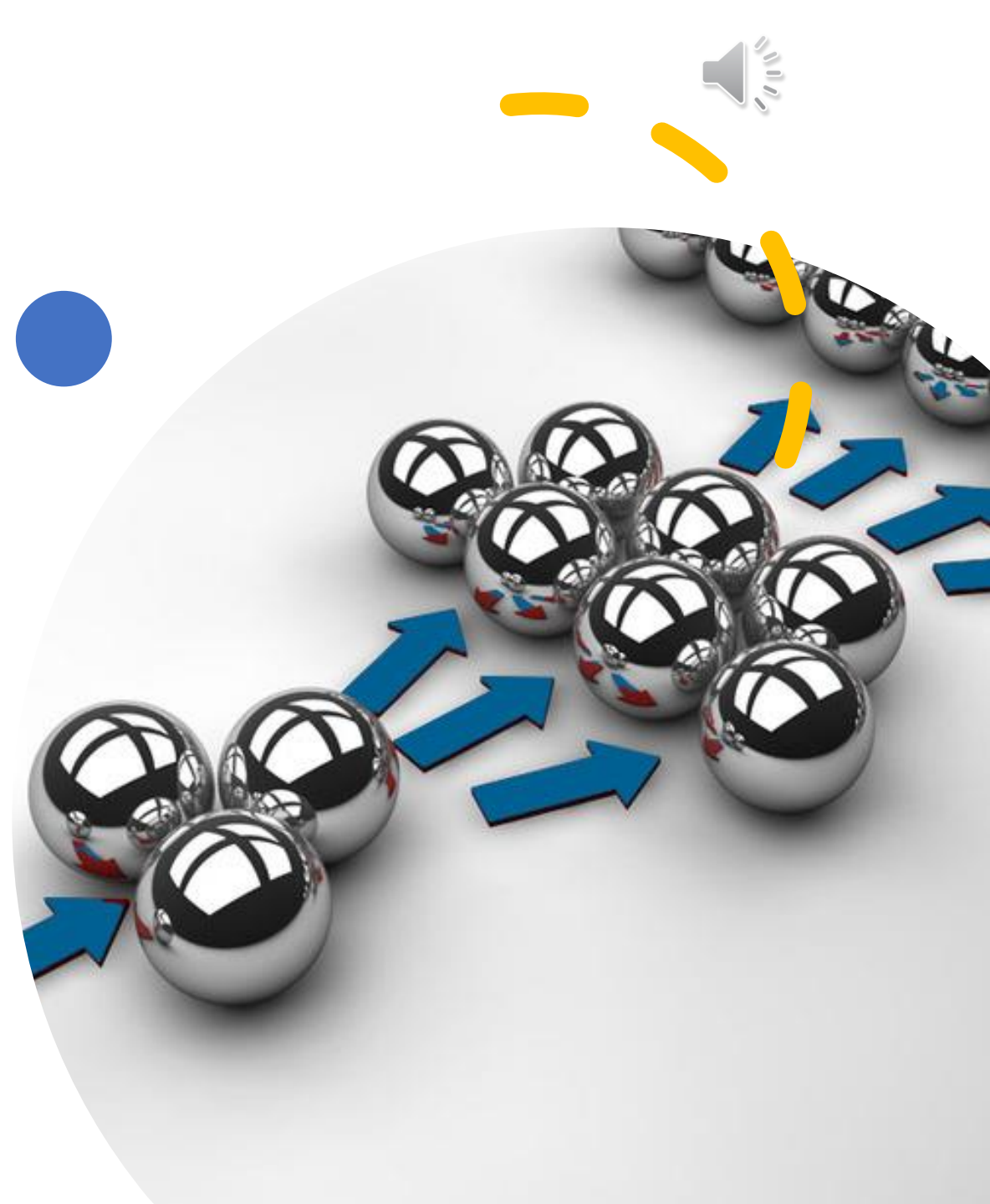
Reference number: 618596-EPP-1-2020-1-SE-EPPKA2-CBHE-JP

This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.



# Impact

- Impact has multiple meanings. It can relate to the total effects of interventions or to longer-term effects on societies, communities, or systems.
- In log frame analysis, it is used with the word result to describe short- and medium-term effects on groups' or individuals' attitudes, abilities, knowledge, or behavior.
- In other cases, the result refers to an intervention's total effects, like impact.





# Impact use

- When paired with the outcome, the impact refers to long-term effects. It also means a development intervention's full impact. It has expected and unforeseen, good and bad, short-term and long-term effects on people, organizations, society, and the environment.
- Effectiveness and broad impact studies overlap. It differs from such a study in two respects. First, impact studies solely analyze effects, while effectiveness assessments consider outputs and impacts. Second, effectiveness studies focus on planned short- or medium-term positive benefits, while impact studies evaluate all consequences, including unforeseen, long-term, and non-target group effects.





# Impact



- Impact is not a criterion like effectiveness or efficiency.
- Effectiveness and efficiency are normative criteria, but the impact criterion recommends considering all relevant repercussions of an action, positive and bad.
- To identify positive and bad outcomes, we need normative criteria like improved well-being of primary stakeholders, poverty reduction utility, etc.

# Types of impact



- As shown below, a study of impact deals with four effect types.
- On the left are the intervention's expected beneficial results and its anticipated negative impacts, which those responsible for it accept as affordable costs.
- The right-hand column contains the unexpected effects, positive as well as negative.

EXPECTED POSITIVE	UNEXPECTED POSITIVE
EXPECTED NEGATIVE	UNEXPECTED NEGATIVE



# Responsibilities and unbiasedness in evaluation



Evaluations should be assessed positively and negatively.



Development intervention planners sometimes overlook or minimize negative impacts.



Ignorance, poor planning, and wishful thinking generate unexpected negative results.



Development initiatives are usually heavily consulted with stakeholders and outside experts, yet things may not go as anticipated.



Sometimes the effects are the opposite of what was planned. Unintended outcomes and their sources are one of evaluation's key jobs.



# Measuring change and determining correlation



When studying impact we face the same technical problems of measuring change and inferring causality as in studies of effectiveness.



However, the first of the two tasks, of measuring change, can be rather more difficult in impact studies than in assessments of effectiveness.



change cannot be measured without a baseline describing the situation before the intervention.



# What is a baseline study?

- A baseline study describes local conditions before a development intervention.
- A baseline study establishes benchmarks for assessing change and progress. Assessments of effectiveness and impact are impossible without baseline information.
- The scope and focus of a baseline study reflect the intervention's objective and the anticipated requirement for data for monitoring and assessment.



# Measuring change and determining correlation



- baseline information collected before the intervention is likely to refer to those aspects of the situation that are expected to change as a result of the intervention, but it may not be relevant for measuring unplanned change.
- To assess unexpected change, we may need an baseline. However, establishing a baseline after the fact is often impossible.
- The second key goal is to determine, with as much certainty as feasible, whether the changes that have occurred since the intervention began were caused by the intervention or would have occurred anyhow. Impact is the difference between actual changes and those that would have occurred without the intervention.
- Counterfactuals are hypothetical patterns against which we compare real changes.



# Measuring change and determining correlation



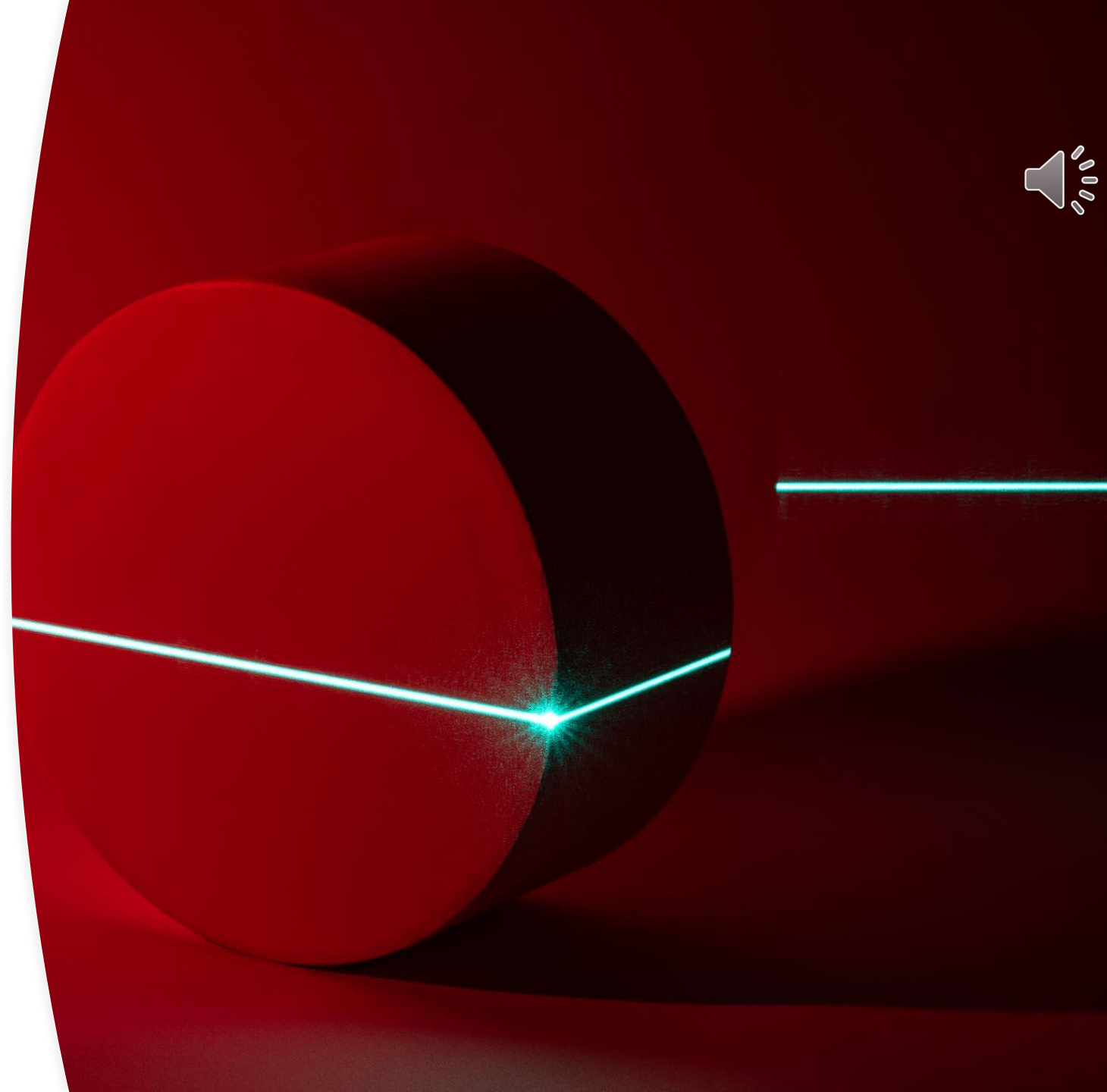
- It is occasionally possible to estimate how the target group would have behaved without the intervention using control groups.
- When the counterfactual cannot be estimated, like in development cooperation, impact statements are weaker.
- The intervention is generally blamed for the alterations if professional knowledge supports it and there is no other reason. This form of reasoning can be sufficient for evaluation; however it is usually less convincing than a control group explanation.
- A textbook on evaluative research methodologies should cover causal attribution in respect to diverse research designs.



# Types of negative unintended effects

---

- Targeting errors
- Substitution and displacement
- Recoil effects
- Fungibility
- Perverse effects



# Types of negative unintended effects



- **TARGETING ERRORS-** Type I and II errors are the main categories.
  - The first type is coverage: Have all subgroups of the targeted category received benefits from the intervention, or has a coverage bias omitted women, children, the old, and the disabled?
  - The second type involves benefit leakage to non-targeted groups. Thus, the question is whether outsiders have taken target group benefits. A problem of overinclusion?
- **SUBSTITUTION AND DISPLACEMENT** - In both circumstances, a target group benefits at the expense of another equally worthy group or category. Subsidized workers replace unsubsidized workers who would have been employed. Displacement would occur if subsidized jobs in one firm reduced workers in other firms.



# Types of negative unintended effects



- **RECOIL EFFECTS**

- These are unforeseen consequences of an intervention on its implementers and managers.
- Development assistance strengthens poor countries and their organizations to handle serious development issues with their own resources.
- In many cases, recipient country organizations are overburdened by externally financed development interventions.
- Evaluations can be used to assess the extent to which development initiatives produce such negative recoil effects.



# Types of negative unintended effects



- **FUNGIBILITY**

- The receiver can transfer limited finances to other activities, some of which might conflict with the donor's objective, by sponsoring an activity. While the donor concentrates on support activities, help may enable unwanted actions.

- **PERVERSE EFFECTS**

- Substitution and displacement, fungibility, as well as some of the recoil effects of aid programmes can all be described as perverse effects.
- Effects referred to by this term are directly opposed to the objectives and goals that the intervention was intended to achieve.



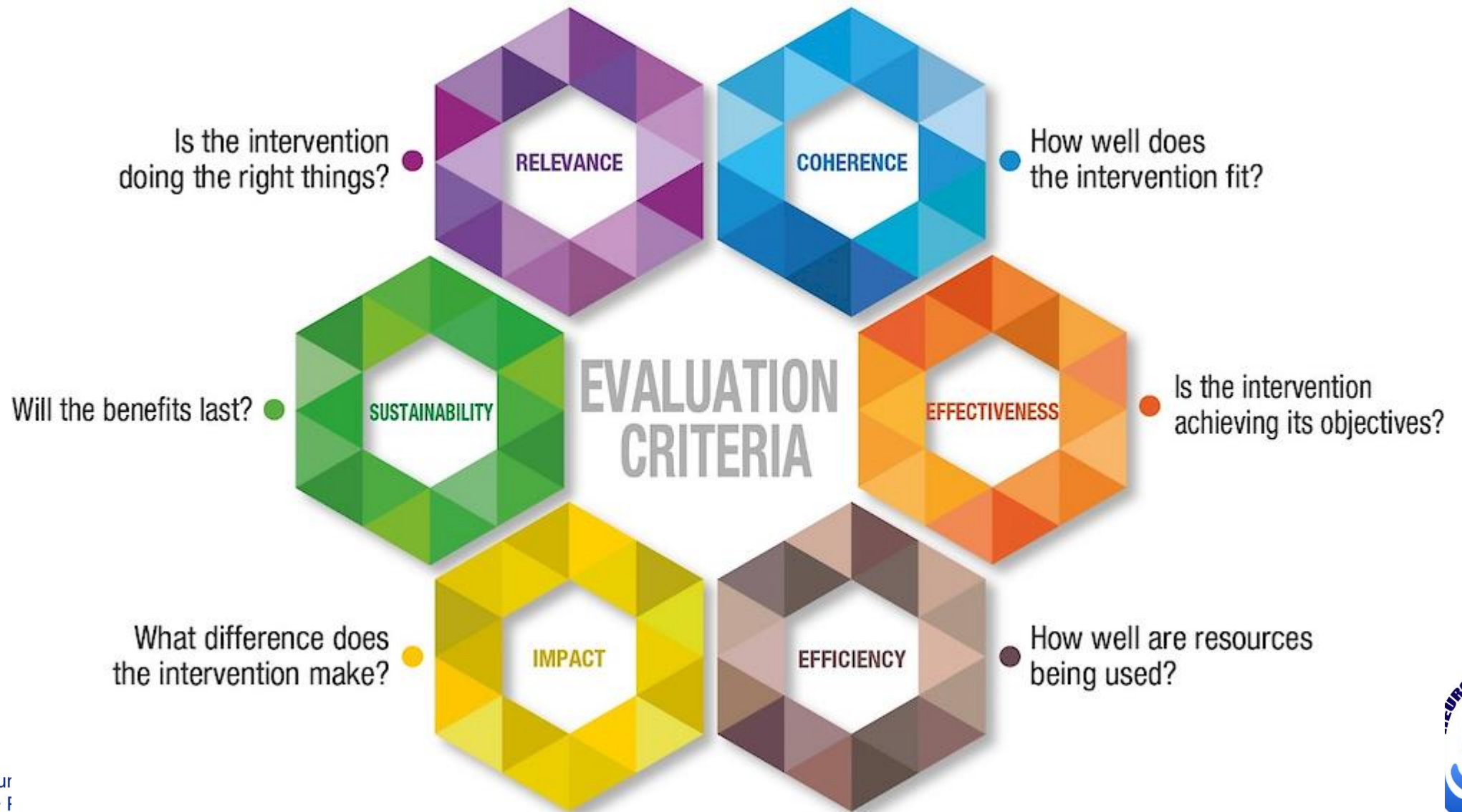
# Standard questions about impact:



- What are the good and negative effects of the intervention on people, institutions, and the environment? How has the intervention affected stakeholder well-being?
- How do beneficiaries and stakeholders estimate the impact of the intervention on themselves?
- What is the impact of the intervention on the managing organization(s) in the recipient country? How much does the intervention increase institutions capacity?
- Can changes throughout the intervention or evaluation period be identified and measured?
- How much of the identified changes can be attributable to the intervention? What would have happened without intervention?
- have alternative explanations for identified alterations been thoroughly investigated and rejected?



# Summary





# Relevance



Co-funded by the  
Erasmus+ Programme  
of the European Union

Reference number: 618596-EPP-1-2020-1-SE-EPPKA2-CBHE-JP  
This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.



# Relevance

- The value and usefulness of the evaluated intervention in the perspectives of key stake- holders.
- Development cooperation intervention is considered relevant if it matches the **needs and priorities of its target group, as well as the policies of partner country governments and donor organizations.**



# Relevance



An intervention is irrelevant if it does not directly or indirectly address major development requirements.

*Although it serves critical needs, this is not enough.*

- Development interventions should align with successful policies and priorities of target groups and others to be meaningful.
- It should also be technically appropriate—an effective, affordable, and side-effect-free remedy.
- The definition implies this. If the intervention doesn't match stakeholder priorities or is technically deficient, it won't fulfil its goals and won't solve the development problem.



# Relevance



- We need to evaluate if the intervention matches end-users' livelihood patterns and social and political conditions and other key stakeholders.
- We must also ensure it matches government policies, administrative systems, and other development groups' operations.
- Important partner country ownership questions.
  - The evaluated intervention: host country autonomy or donor choice adaptation?
  - And how much do host country actors and donor agencies run it?
  - Shared ownership with intended beneficiaries and citizens?
  - Are there accountability and public participation mechanisms?



# Standard questions about relevance:



---

Is the intervention consistent with the livelihood strategies and living conditions of its target group? How urgent is it from the point of view of the target group?

---

Is the intervention consistent with national and regional development policies and administrative systems of the partner country? Is it consistent with partner country ownership policy?

---

Is the intervention a technically adequate solution to the development problem at hand? Does it eliminate the main causes of the problem?

---

Do proposed innovations have a potential for replication?

---





# Sustainability



Co-funded by the  
Erasmus+ Programme  
of the European Union

Reference number: 618596-EPP-1-2020-1-SE-EPPKA2-CBHE-JP  
This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.





# Sustainability

- Sustainability is the possibility that an intervention's benefits will last after the withdrawal of donor support.
- Sustainability is part of the impact criterion, but since it's a recurring topic of great relevance in development cooperation, we can treat it separately.
- Sustainability must be specified for the intervention under review. varied interventions have varied time frames and development functions.

# What determines whether the examined intervention's results be sustained into the future?



**Partner country priorities**



**Partner country ownership and participation**



**Institutional and cultural factors**



**Technological factors**



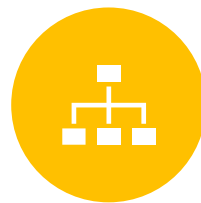
**Exit strategy**



**Environmental factors**



**Financial factors**

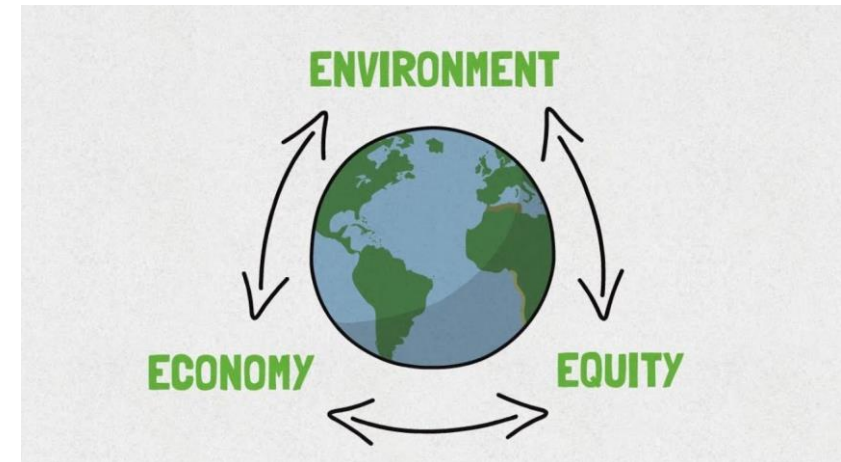


**Management and organisation**

# Standard questions about sustainability:



- Is the intervention consistent with partners' priorities and effective demand? Is it supported by local institutions and well integrated with local social and cultural conditions?
- Are requirements of local ownership satisfied? Did partner country stakeholders participate in the planning and implementation of the intervention?



# Standard questions about sustainability:

---



Are relevant host-country institutions characterized by good governance, including effective management and organization?

---

Is the technology utilized in the intervention appropriate to the economic, educational, and cultural conditions in the partner country?

---

Do partners have the financial capacity to maintain the benefits from the intervention when donor support has been withdrawn?

---

Is the intervention compatible with a sustainable use of natural resources? Or is it harmful to the natural environment?

---





# Efficiency



Co-funded by the  
Erasmus+ Programme  
of the European Union

Reference number: 618596-EPP-1-2020-1-SE-EPPKA2-CBHE-JP

This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.



# Efficiency



Efficiency is the ratio of the value of the results of an intervention to the value of the resources used to produce them. An intervention is technically efficient when its productive capacity is fully utilized.



An intervention is optimally efficient if its value is greater than the value of any alternative use of these resources.

*If the same resources could have produced better results in some other way, or if the same results could have been produced with fewer resources, it is less than fully efficient.*



Economists distinguish between different types of efficiency.

*Technical efficiency - simplest, efficiency*  
*Allocative efficiency- more complex*  
*concept of efficiency*



Co-funded by the  
Erasmus+ Programme  
of the European Union

Reference number: 618596-EPP-1-2020-1-SE-EPPKA2-CBHE-JP  
This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.



# Technical efficiency vs. allocative efficiency



• Technical efficiency	• Allocative efficiency
<ul style="list-style-type: none"><li>• Intervention is evaluated against other ways of achieving the same concrete objective, regardless of the value of that objective.</li></ul>	<ul style="list-style-type: none"><li>• In an assessment of allocative efficiency, by contrast, an intervention is evaluated against every alternative use of the same resources, at least in principle.</li></ul>
<ul style="list-style-type: none"><li>• Assessment of technical efficiency focuses on the relation between inputs and outcomes (or outputs) and takes the objective as given</li></ul>	<ul style="list-style-type: none"><li>• An assessment of allocative efficiency raises the more far-reaching question of whether the intervention is economically worthwhile, given the alternatives foregone.</li></ul>



# Economic evaluations



Assessments of efficiency are known as *economic evaluations*. By a standard definition, an economic evaluation is a comparative analysis of alternative courses of action in terms of both their costs and their consequences.



The object of comparison is usually another intervention, but it can also be a constructed model of best practice or standard of performance.



Among the full economic evaluations, the following are the most important: cost-effectiveness analysis (CEA), cost-utility analysis (CUA) and cost-benefit analysis (CBA).

# Economic evaluations



Healthy, attractive ecosystems and human life have costs and rewards. After the CBA, non-monetary costs and benefits are sometimes considered separately.



The final analysis compares costs and benefits. Otherwise, the intervention is not worthwhile until benefits exceed costs.



CBA evaluates programmes, policies, and actions. Helping limited assets be used economically is the goal. Methodically comparing pros and cons is one of its features.



Distributive effects analysis is another drawback. CBA deems an intervention socially useful and economically worthwhile if it increases society's total happiness as measured by the market or a surrogate like consumers' "willingness to pay."



Calling this a CBA limitation may be misleading. The CBA prioritizes economic efficiency over equity and justice.

# Standard questions about efficiency:



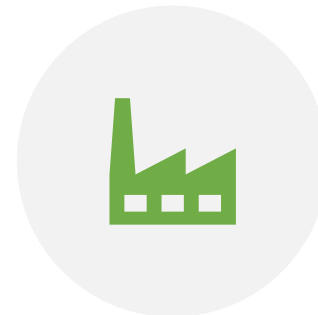
Has the evaluated intervention been managed with reasonable regard for efficiency?



What measures have been taken during planning and implementation to ensure that resources are efficiently used?



Could the intervention have been implemented with fewer resources without reducing the quality and quantity of the results?



Could more of the same result have been produced with the same resources?

# Standard questions about efficiency:



Could an altogether different type of intervention have solved the same development problem but at a lower cost?



Was the intervention economically worthwhile, given possible alternative uses of the available resources? Should the resources allocated to the intervention have been used for another, more worthwhile, purpose?



Co-funded by the  
Erasmus+ Programme  
of the European Union

Reference number: 618596-EPP-1-2020-1-SE-EPPKA2-CBHE-JP  
This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.



# Reference

Looking Back, Moving Forward - Sida Evaluation Manual

*(Molund, S., Schill, G., 2004. Looking back, moving forward: Sida evaluation manual. Sida, Stockholm).*



**Thank you**



Co-funded by the  
Erasmus+ Programme  
of the European Union

Reference number: 618596-EPP-1-2020-1-SE-EPPKA2-CBHE-JP  
This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.

