



**IEO** | INDEPENDENT  
Evaluation Office  
United Nations Development Programme



وزارة التخطيط والمتابعة والإصلاح الإداري  
Ministry of Planning, Monitoring  
and Administrative Reform



مصر  
2030



## LEAVING NO ONE BEHIND: EVALUATION for 2030

2019 National Evaluation  
Capacities Conference

BETTER CRITERIA FOR BETTER EVALUATION  
USING EVALUATION CRITERIA IN A NATIONAL  
EVALUATION CONTEXT

MEGAN KENNEDY-CHOUANE &  
VELAYUTHAN SIVAGNANASOTHY

 @mkennedychouane  
@OECD\_EVALNET  
<https://oe.cd/criteria>

#BetterEvaluation  
#NECdev  
#EvalCriteria

# Purpose of today's training



*Our goal is to support you in producing or using evaluations, to improve evaluation quality and uptake in national sustainable development policies/programmes...*

*...by sharing new criteria definitions, lessons from Sri Lanka and insights from the OECD/DAC consultation process on revising evaluation criteria.*

# Agenda

09:00 – 9:30	Introductions & context
9:30 – 10:40	Key evaluation concepts & principles National evaluation systems (policy, set-up) Discussion
10:40 – 11:00	Coffee/tea break
11:00 – 13:00	Presentation of criteria update process and new definitions (30 minutes) Contextualization and adaption in national context (30 min) Discussion
13:00 – 14:00	Lunch break
14:00 – 15:00	Case study example
15:00 – 15:30	Discussion of challenges
15:40 – 16:00	Coffee/tea break
16:00 – 17:20	Addressing challenges, principles for use and your projects
17:20 – 17:30	Wrap up and evaluations

# Logistics

- Training mentality!
- On your table: fidget toys, sticky notes, markers
- Emergency exits
- Water and restrooms

# Setting your own agenda: On a sticky note

**Pink/Orange:** What motivates you?  
Why you are here today 😊

**Yellow/Green:** What is one problem you have or issue you face in your work, that you would like to work on/address today?



# Introduce yourself to the person next to you - then present that person to the class

- Name + What your name means
- Two places: where born and where live
- Institution/country and role in evaluation
- One thing you have in common (can't be related to this conference!)

Present to class:

- Your names & roles
- One thing same



# Trainers



- Megan Kennedy-Chouane
- Santa Cruz, USA & Paris, France
- OECD
- Advise on evaluation policy and practice, support collaboration and learning in evaluation (incl. capacity development)
- Both studied management

- Velayuthan Sivagnanasothy
- Colombo, Sri Lanka
- Ministry of National, Policies, Economic Affairs, Resettlement and Rehabilitation, Northern Province Development and Youth Affairs, Sri Lanka

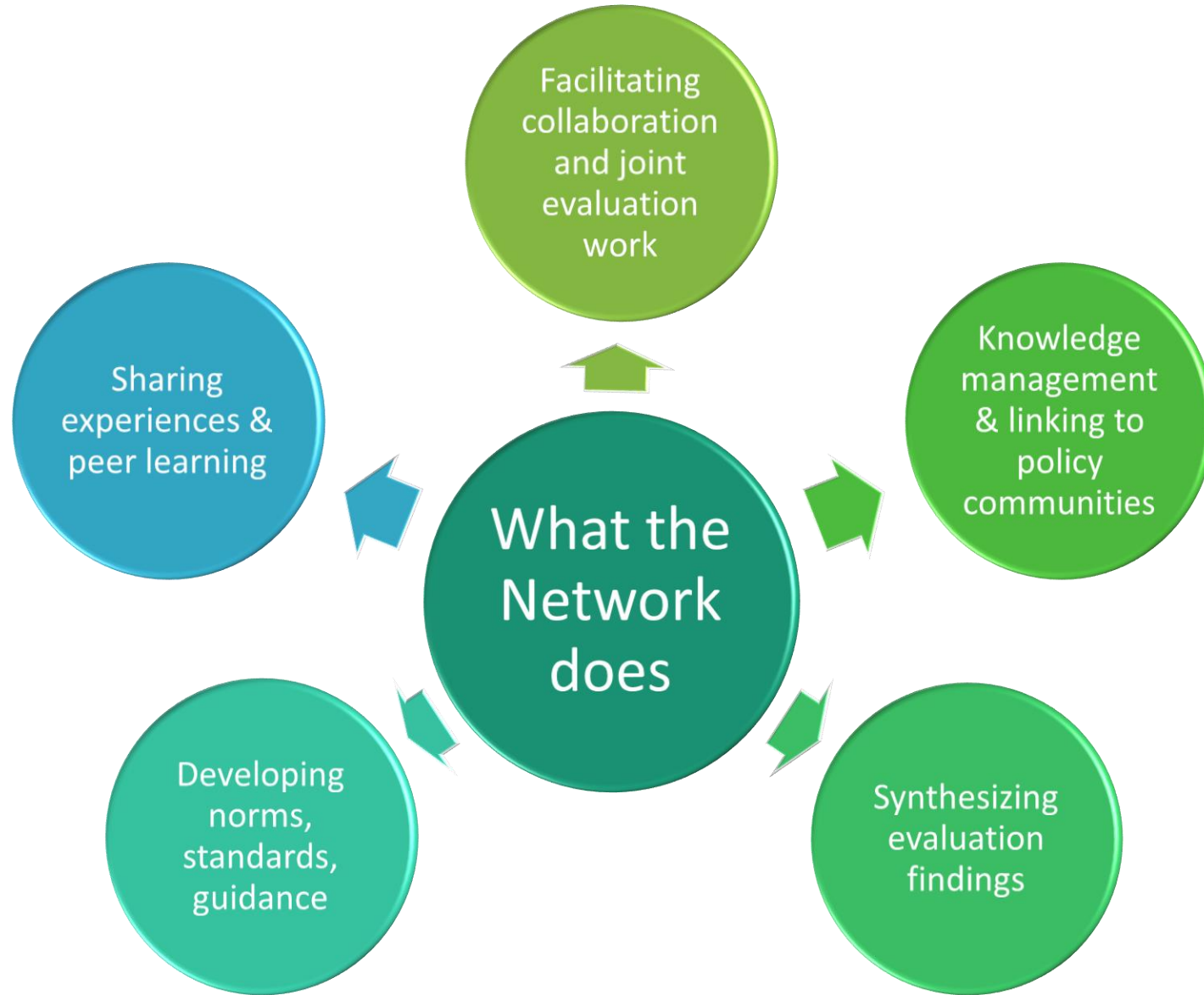
SETTING THE STAGE: BASIC CONCEPTS  
(20 MIN)



# Who is EvalNet?

- Network in the OECD Development Assistance Committee (DAC)
- Bringing together evaluation heads and managers of 30 OECD Ministries of Foreign Affairs as well as development agencies, five regional development banks, the World Bank, IMF and UNDP
- Providing a forum for dialogue, exchange and knowledge sharing
- The current Chair of the Network is Per Bastøe (Norway).
- Supported by Secretariat within the Development Co-operation Directorate of the OECD.





# Difference between M and E

**Monitoring** and **Evaluation** are generally viewed as distinct but complementary functions. Both are needed to better manage and account for results.

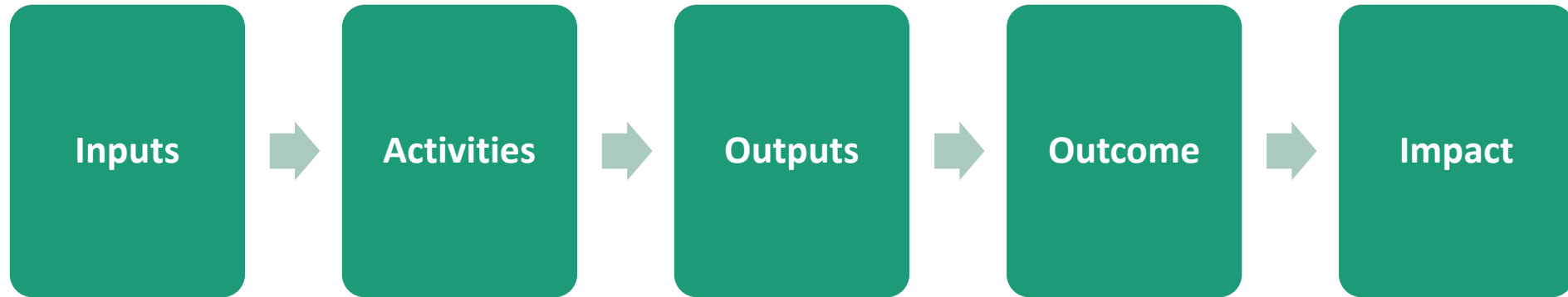
Monitoring	Evaluation
<ul style="list-style-type: none"><li>• Periodic, using data routinely gathered, generally internal</li><li>• Tracks progress against a small number of target indicators</li><li>• Usually quantitative</li><li>• Assumes appropriateness of project, activities, objectives and indicators</li><li>• Cannot indicate causality</li><li>• Reports on progress to managers and alerts them to problems requiring attention and action</li></ul>	<ul style="list-style-type: none"><li>• Generally episodic, in-depth, often external</li><li>• Can question the rationale and relevance of the project and its objectives</li><li>• Can address why and how intended results were or were not achieved</li><li>• Can explore unintended results and effects</li><li>• Can address attribution, cause-effect, linkage or contribution</li><li>• Can provide guidance and recommendations for future directions</li><li>• Can use data from different sources and a variety of methods</li></ul>

## The power of evaluating results

- If you do not measure results, you can not tell success from failure.
- If you can not see success, you can not learn from it.
- If you can not recognise failure, you can not correct it.
- If you can demonstrate results, you can win public support.

*Adapted from Osborne & Gaebler, 1992 and R. Rist, 2000*

# Results chain



**Monitoring:** What has been invested, done and how are we progressing towards the achievement of objectives?

**Evaluation:** What has been achieved, how, why and for whom?

# Defining evaluation

- **Development evaluation** is the systematic and objective assessment of a development programme or policy, its implementation and its results.
- Evaluation is a key tool in efforts to improve **accountability and performance** of development programmes and policies.
- The role of evaluation is to support **better sustainable development results** (by supporting effective development policies, programmes and co-operation)



# Why evaluate?

## Evaluation for **use!**

```
graph TD; A[Evaluation for use!] --> B[Decision-making]; A --> C[Learning]; A --> D[Accountability];
```

### Decision-making

- Should we change course, close down, scale-up?
- Do we need to change strategy?

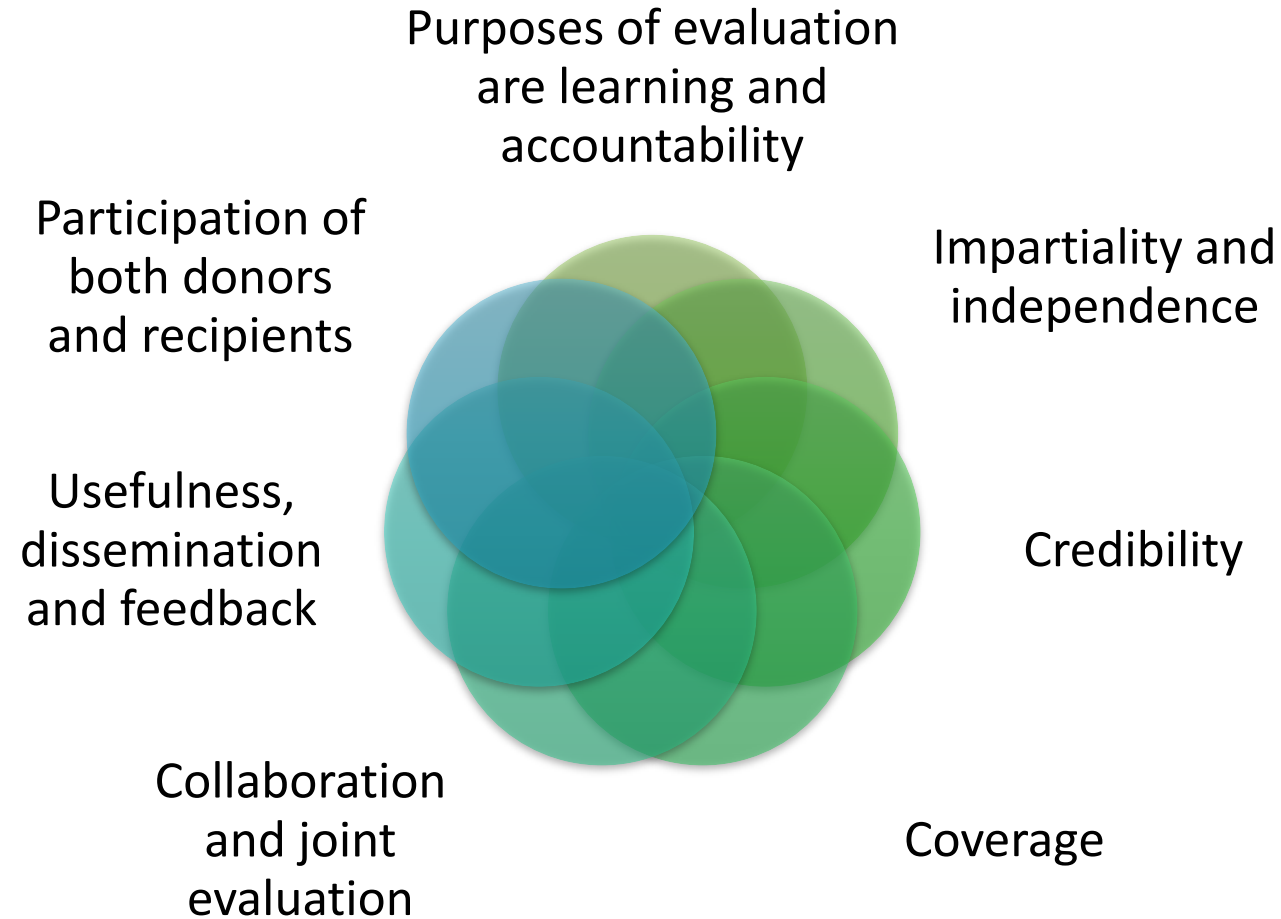
### Learning

- What worked, what didn't and why
- How to achieve desired results
- Understanding development processes and influence of context

### Accountability

- For achieving results
- For use of resources
- To funders, beneficiaries, governments, publics, etc.

# Core principles of evaluation



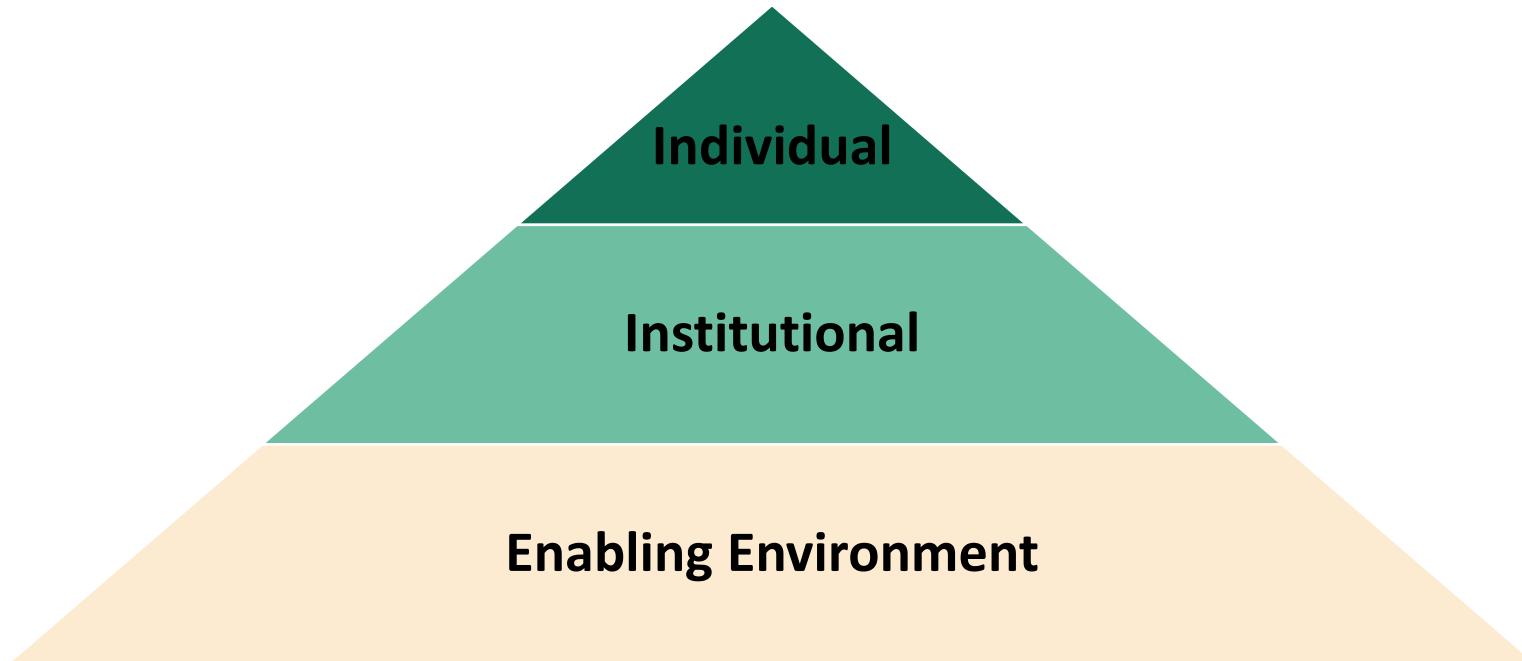


# The importance of credibility and impartiality for a good evaluation system

- Evaluation process should be independent from the process concerned with policy-making and the delivery.
- Credible evidence vs. selling to the public
- Ways to strengthen independence:
  - Semi-independent evaluation unit, direct reporting to head or evaluation committee (depends on organisational structure)
  - Control over own budget and work programme
  - Reporting lines
  - Use of consultants
- Criteria are part of a credible evaluation system

# Evaluation capacity

- Evaluation capacity is not *just about technical skills*, but involves strengthening *institutions and systems*, and developing an *enabling environment* for evidence-based policy making (supply & demand).
- Evaluation capacity development (ECD) is defined as: “unleashing, strengthening and maintaining capacities for evaluation,” at three levels:



# Steps towards establishing an evaluation system

- Reflect on and define the purpose and need for M&E:
  - What do you see as the primary driver(s) in demand for evaluation in your national context today?
- Find an appropriate role for M&E within your system
- Develop an evaluation policy or mandate, in line with the legal and institutional frameworks
- Identify useful evaluation manuals, guidelines, tools

# National Evaluation Systems

Examples from Sri Lanka and other countries

Reactions and thoughts?



REFRESHMENTS BREAK

EVALUATION CRITERIA  
CONCEPTS (30 MIN)

# What are evaluation criteria?

- Broad guides to help us ask the right questions – focus on results instead of (input and activities)
- Relevance, effectiveness, efficiency, impact and sustainability
- First set out by the OECD/DAC in evaluation principles, defined, have become very widely used – cornerstone of evaluation practice
- But they are just the foundation! Will not solve all problems of evaluation, but they are a contribution

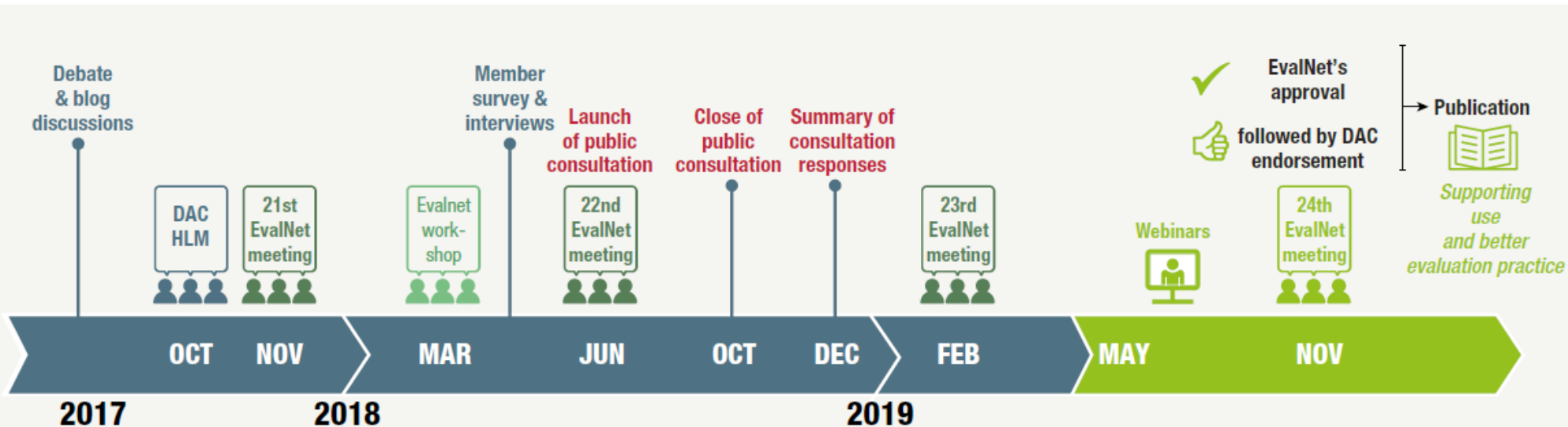




# Why update the criteria?

- 2015: Adoption of Agenda 2030 Sustainable Development Goals & Paris Climate Agreement UNFCCC
- Address common challenges and weaknesses (misinterpretation) to strengthen evaluation practice
- Reflect lessons learned over the years
- Request from the Development Assistance Committee (DAC)

# DEVELOPMENT EVALUATION CRITERIA ADAPTATION TIMELINE



Discussion of experiences with the criteria and adaptation at international evaluation events

Drafting & Consultations on text

Developing Guidance

**Global consultations:** The adaptation process has included a major global consultation, involving hundreds of stakeholders from many backgrounds and countries.

# Main challenge: use

- Applied simplistically or mechanistically
- Donor-imposed or made mandatory
- Covering too many questions in a shallow way (all criteria)
- Not sufficiently contextualized
- Weak coverage of equity issues
- Though words are widely use, many are not fully familiar with definitions and/or intended role of criteria

## Views 1a: Perceived strengths of the set of criteria *(Ten most frequent responses)*

Perceived Strengths
<i>Universal acceptance &amp; use across evaluation profession &amp; beyond</i> – creation of a common language & understanding; a normative framework
<i>Standardisation/consistency</i> – provide scope for comparability/synthesis across evaluands/contexts
<i>Comprehensiveness/completeness</i> – coverage of key areas required for accountability and learning
<i>Simplicity and clarity</i> – readily understandable, clear formulation
<i>Neutrality</i> - acceptable across cultures/political contexts
<i>Universality</i> - applicable to different evaluands; institution types; policy/intervention areas etc
<i>Utility</i> – provide relevant information to support improvement/change
<i>Results focus</i> – emphasise the importance of results at different levels
<i>Adaptability/flexibility</i> - can be tailored for different evaluands, contexts etc
<i>Conciseness/feasibility</i> – Limited in number so realistic to implement

### Illustrative quotes

*“The DAC criteria are a very useful elaboration of the merit, worth and value trilogy and have served the evaluation community well.”*

*‘Without them, evaluation of development interventions would become quite ad hoc in terms of performance assessment.’*

*‘They are clear, concise and internationally respected.’*

*‘If they didn’t exist, they would have to be invented.’*

## Views 1b: Perceived weaknesses of the set of criteria (Ten most frequent responses)

### Perceived Weaknesses

**Insufficient encompassing of SDG agenda** – Including issues such as complexity, inclusiveness ('no one left behind') and partnerships

**Limited applicability in to different evaluation types** - Mostly applicable to project/programme evaluations, rather than strategy/policy/institutional evaluations etc

**Linearity** – Do not encompass systems thinking, interconnectedness

**Restricted in scope** – Do not encompass all types of development assistance now being applied

**Limited measurability** – In aggregate, do not permit robust measurement across all criteria

**Insufficient recognition of context change/adaptive capacity** – Provide a largely 'static' picture

**Narrow approach to 'results' rather than 'change'** - Do not emphasise transformative change

**Insufficient focus on gender, equity, human rights concerns** – Not explicitly integrated or prioritised

**Vague** – Lack adequate specification as a set or as individual criteria

**Promote summative judgements** – Rather than prioritising transformation, improvement or learning

### Illustrative quotes

*"They are a good servant but a lousy master."*

*'They should include the following concepts: systemic approaches, emergence, integration through multi-dimensional policies and partnerships, inclusion leaving 'no one behind', cultural beliefs, social norms.'*

*'They are not very useful for corporate and strategy evaluations; the political economy analysis dimensions; and [nor] do they respond to the complexity of the SDGs.'*

## Views 4: Retention, adaptation, removal of current criteria

### Overall views *(most frequent responses)*:

- **Better to adapt existing criteria** than to introduce a full new set
- **Strong justification** would be needed for individual criteria removal

### Adaptation:

- **Refinement required** rather than 'full' or 'wholesale' adaptation
- Adapt to include **recognition of SDGs and complexity**
- Specify **inter-relationships across criteria**
- Need to include more explicitly **gender, human rights, equity concerns**

### Guidance

- **Stronger guidance** required for implementation
- **Include examples** in guidance

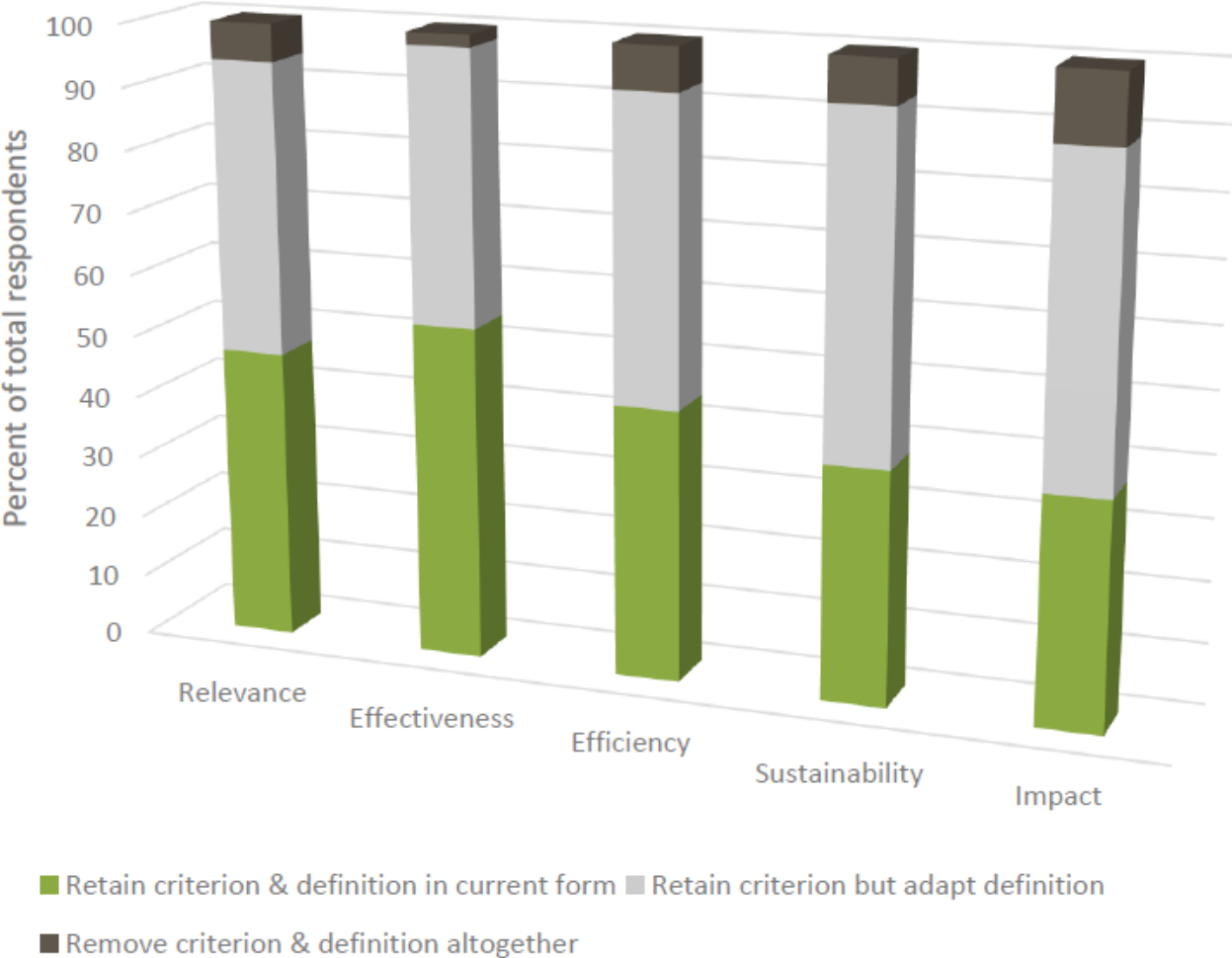
### Illustrative quotes

- *Overall the criteria have proved their usefulness over time and don't need to be substantially changed.*
- *'Don't fix what isn't broken'*
- *'Revision, not reform!'*
- *'Not sure what is the purpose of changing them. It will just create discontinuity.'*
- *'I'm not usually conservative, but I would really think twice about changing a winning team.'*
- *'One should try to invest in reforming the catalogue without transforming it into a new set of criteria. Sharpening the criteria is the most important aspect.'*
- *Change as much needed but as little as possible*



# Views 4: Current criteria - Retain, adapt or remove?

Answered: 466



**High majority of respondents** preferred to *Retain criterion & definition in full or Retain the criterion but adapt definition* (over 89% of responses for all five criteria)

- **Relevance:** 94% of respondents prefer to *Retain criterion & definition in current form* (47%) or *Retain criterion but adapt definition* (47%)
- **Effectiveness:** 98% prefer to *Retain criterion & definition in current form* (54%) or *Retain criterion but adapt definition* (44%)
- **Efficiency, Sustainability** 93% prefer to *Retain criteria & definition in current form* (44%, 38%) or *Retain criteria but adapt definitions* (49%, 55%)
- **Impact:** 89% prefer to *Retain criterion & definition in current form* (37%) or *Retain criterion but adapt definition* (52%). 11% preferred to 'Remove' the criterion.

## Views 7: What needs to change?

**Headline themes** from consultation: (most frequent responses)

- Greater (more explicit) attunement with **SDG narrative** e.g. societal benefits/power/equity/inclusiveness
- Recognise **complexity/systems models**
- Specify/emphasise **interconnectedness** – the criteria stand ‘in relation’ not ‘in isolation’
- Ensure applicability to **policy, programme, systems, institutional and strategic evaluations**
- Ensure strong presence of **gender, equity, human rights throughout**
- Support implementation through **stronger guidance**

### Illustrative quotes

*‘The main challenge is to develop the set of criteria as a guide, not as a religion!’*

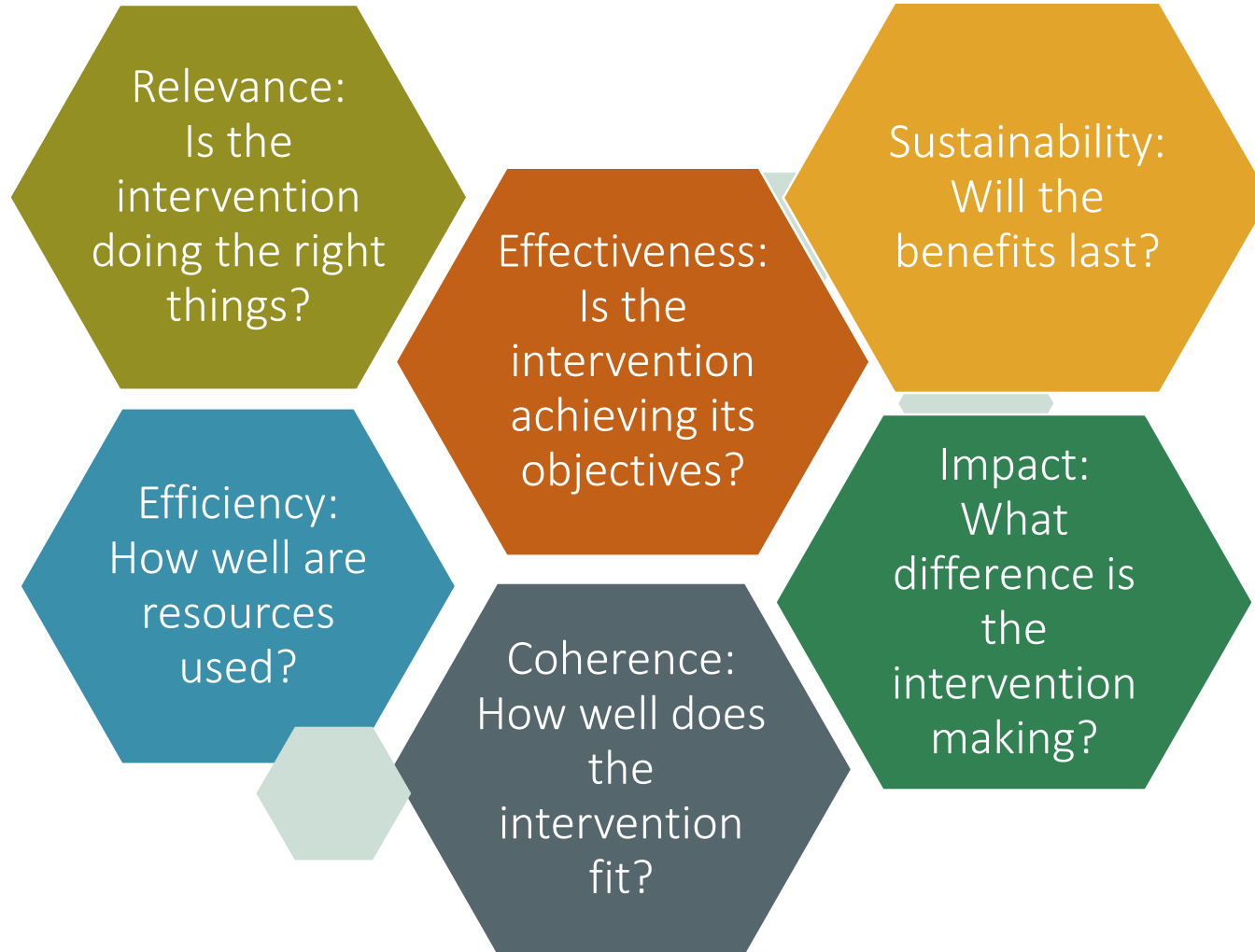
*‘They are a framework that needs to be supplemented with specific questions - but the framework is strong and quite embedded internationally.’*

*‘They have stood the test of time and, with adaptation, will continue to do so in future.’*



# EVALUATION CRITERIA

DRAFT FOR CONSULTATION – NOT FOR CITATION



# New criteria key features

- New and improved definitions
- Retaining conceptual clarity and keeping the definitions as simple as possible.
- One major new criterion – Coherence – to better capture synergies, linkages, partnership dynamics, and complexity.
- Supporting use and addressing confusion:
  - an introduction on the intended purpose of the criteria;
  - guiding principles for use; and
  - an accompanying guidance (forthcoming).

# Main features (continued)

- Reflecting the integrated nature of sustainable development and current policy priorities
- Promote a more interconnected approach to the criteria, including examination of synergies and trade-offs.
- Yet high-level enough to ensure they will remain relevant as policy priorities and goals change.
- Ensuring applicability across diverse interventions (beyond projects)

Each criteria is a lens, giving a different perspective of the intervention implementation and results...



Together, they provide a more complete picture.

ARE WE DOING THE RIGHT THINGS?

**RELEVANCE:** The extent to which the intervention objectives and design respond to affected people's, global, country, and partner/institution needs, policies, and priorities, and remain so if circumstances change.

Note: "Respond to" means that the objectives and design of the intervention are sensitive to the economic, environmental, social, equity, political economy, and capacity conditions in which it takes place. "Partner/institution" includes government (national, regional, local), civil society organisations, private entities and international bodies involved in funding, implementing and/or overseeing the intervention. Relevance assessment involves looking at differences and trade-offs between different priorities or needs. It requires analysing any changes in the context to assess the extent to which the intervention can be (or has been) adapted to remain relevant.

# Explanation for the definition

- Need to “Raise the bar” on evaluating relevance (donor and country policies is far too broad – rare that something would be implemented and not be relevant to something in policies)
- By separating out Coherence, we clarify both concepts in order to deepen evaluation analysis.
- Adding ‘design’ and ‘implementation’ to capture other elements related to the quality of the intervention (not only objectives)
- Making time dimension more explicit: While relevance is generally evaluated looking at the point of time before the intervention starts, the ongoing relevance should also be evaluated. This is now included in the definition
- Highlight on the groups that may be excluded from the identification of priorities/policies.

HOW WELL DOES THE INTERVENTION FIT?



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

The extent to which other interventions support or undermine the intervention, and vice versa. Includes internal coherence and external coherence:

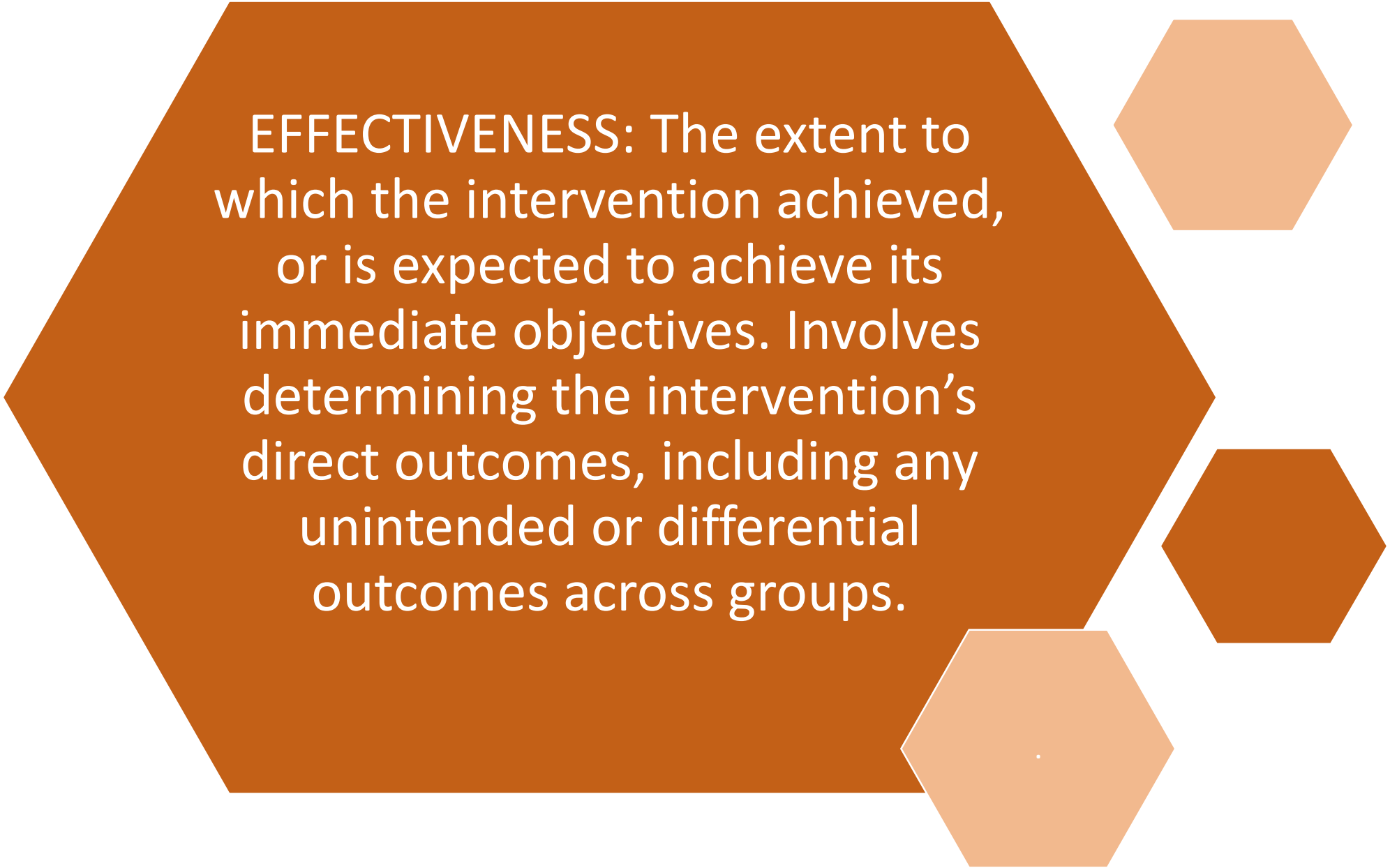
**Internal coherence** addresses the synergies and interlinkages between the intervention and other interventions carried out by the same institution/government, and consistency of the intervention with the relevant international norms and standards to which that institution/government adheres.

**External coherence** considers the consistency of the intervention with other actors' interventions in the same context. Includes complementarity, harmonisation and co-ordination with others, and the extent to which the intervention is adding value while avoiding duplication of effort.

# EXPLANATION OF THE DEFINITION

- Capturing a perspective not covered previously
- Promoting understanding of the role of an intervention within a particular system
- Responding to the need for greater attention to coherence
- Encouraging analysis of consistency with commitments under international law or agreements

IS IT WORKING?



**EFFECTIVENESS:** The extent to which the intervention achieved, or is expected to achieve its immediate objectives. Involves determining the intervention's direct outcomes, including any unintended or differential outcomes across groups.

# EXPLANATION OF DEFINITION

- This is a simple criteria, keep it clear and simple
- Concerned with more closely attributable results distinct from impact
- Enquiring about immediate unintended results and the distribution of results across different groups to encourage more thoughtful analysis, attention to equity

ARE RESOURCES BEING USED WELL?



**EFFICIENCY:** 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 outputs, outcomes and impacts, in the most cost-effective way possible, as compared to feasible alternatives in the context.

“Timely delivery” is within the intended timeframe, or a timeframe reasonably adjusted to the demands of the evolving context.

It may include assessing operational efficiency (how well the intervention was managed).

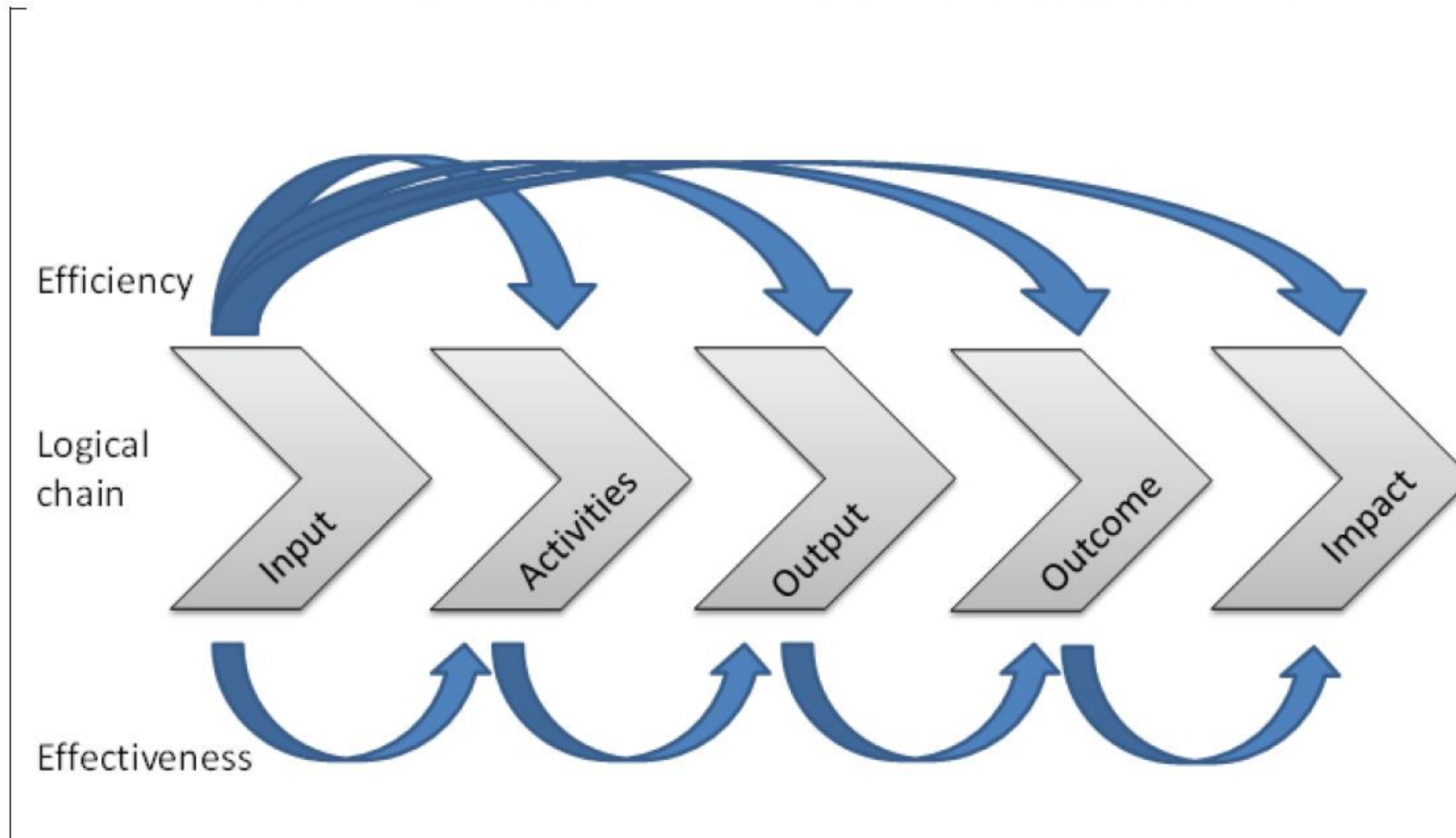


# EXPLANATION OF THE DEFINITION

- Looking at inputs relative to the entire results chain (but also allowing flexibility to focus analysis on the most relevant part of the results chain)
- Valid cost-benefit analysis requires comparing the value of the intervention with pertinent counterfactuals
- Reference to operational efficiency (important area, often focus)
- Implementation issues are of great interest to evaluation stakeholders.
- (competing understandings)



**Figure 1 Efficiency and effectiveness in the logical framework**



<sup>1</sup> Contact addresses: [robrecht.renard@uantwerpen.be](mailto:robrecht.renard@uantwerpen.be) and [slister@mokoro.co.uk](mailto:slister@mokoro.co.uk)

IS THE INTERVENTION  
MAKING A DIFFERENCE?

**IMPACT:** 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 ultimate significance and potentially transformative effects of the intervention.

Beyond the primary and immediate effects, this criterion seeks to capture the indirect, secondary and potential consequences of the intervention. It does so by examining the holistic and enduring changes in systems or norms, and potential effects on people's well-being, human rights, gender equality, and the environment.

It seeks to identify the higher plane social, environmental and economic effects of the intervention especially those that are longer term or broader in scope than those already captured under the effectiveness criterion.

# EXPLANATION OF DEFINITION

- Emphasis on the significance and transformational nature of the effects
- Not a narrow (attribution-focused) definition
- Common understanding of the term and its meaning
- Distinguish between Effectiveness and Impact



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

Includes an examination of the financial, economic, social, environmental, and institutional, capacities of the systems needed to sustain net benefits over time.

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.

Involves analyses of resilience, risks and potential trade-offs.

# EXPLANATION

- Original definition too donor centric and focusing only on external funding
- Current definition showing that sustainability has various dimensions
- Encouraging analysis of potential trade-offs, and of the resilience of capacities/systems underlying the continuation of benefits
- Retaining the term “net benefits” to focus on the overall value, taking into account the costs of producing and maintaining benefits

# Conclusions & Next Steps

- Far reaching feedback
- On criteria – used to develop new definitions and guidance.
- In addition, developing plans for new work focused to address other weaknesses in evaluation practice and systems.
- Ex. Blended finance





Building the rest of  
the house:  
Localization and  
adaption of  
OECD/DAC criteria  
in a national  
context



# Conclusion

- Adaption of international norms , standards and criteria strengthens national evaluation systems
  - Common terminology
  - Harmonization
  - Comparability across countries and donors
  - Improves quality, helps rating , sound evidences for decision making
  - Professionalization of evaluation
  - Helps joint collaborative evaluation
- 2. Localize, contextualize and adapt norms, standards and criteria (not to consider as a straight jacket or mechanical application but adapt flexibly

# Reactions and thoughts?

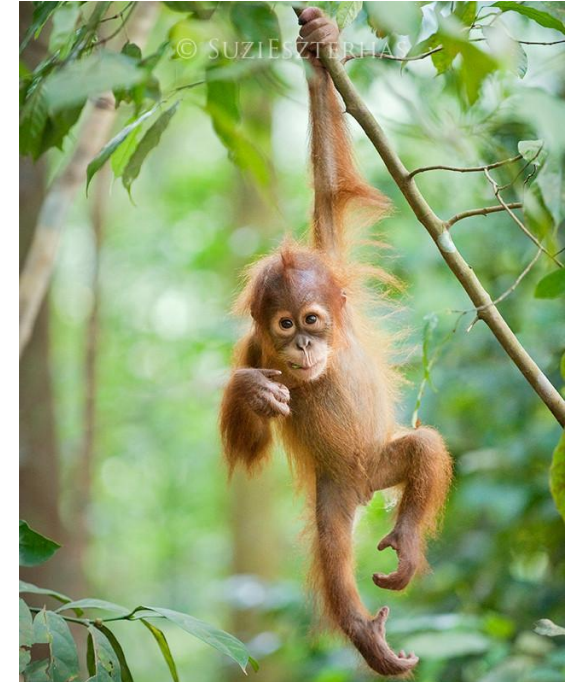
Please discuss at your table for 10 minutes  
– then report back on key 3-4 issues

LUNCH BREAK: RETURN AT 14:00

# PRINCIPLES FOR USE AND FACING COMMON CHALLENGES IN COUNTRIES

# Evaluation case study: Fisheries

- Read the case study (10 minutes)
- Describe theory of change (15 minutes)
- World Café Flip charts
  1. Develop evaluation questions
  2. Refine the evaluation questions, improve and finalize 2 questions
  3. Let's say you have answered the question – what do you do with the answer? Who will use the information



# Small group discussion

- In your small group, share what do you see as the main challenges and issues in conducting evaluations? In supporting evaluation use?
- Pick one *rapporteur* to share your main ideas (2 minutes max)

BREAK



# REACTIONS AND THOUGHTS ON ADDRESSING CHALLENGES



WRAP UP: CHECK OUT WITH ONE WORD YOU  
ARE TAKING WITH YOU INTO THE REST OF  
THE WEEK

# Thank you!

Please fill out your evaluations!

[megan.kennedy-chouane@oecd.org](mailto:megan.kennedy-chouane@oecd.org)

[sivagnanasothy@hotmail.com](mailto:sivagnanasothy@hotmail.com)

@OECD\_EVALNET criteria

