OECD DAC Evaluation Criteria: Summary of consultation responses

November 2018

Consultation Process

Ran March-end October 2018

Content:

- OECD DAC Network member survey
- Stakeholder consultation (quantitative/qualitative) in three languages
- Interviews with key stakeholders
- Presentation/discussion at international meetings/seminars

Input generated:

- 691 survey responses

Member survey	85
Stakeholder consultation – English	576
Stakeholder consultation – French	21
Stakeholder consultation - Spanish	9
TOTAL	**691**

- Over 700pp of qualitative comments
- 11 separate written submissions

STAKEHOLDER CONSULTATION
adapting the DAC evaluation criteria

Who responded? - Stakeholder consultation
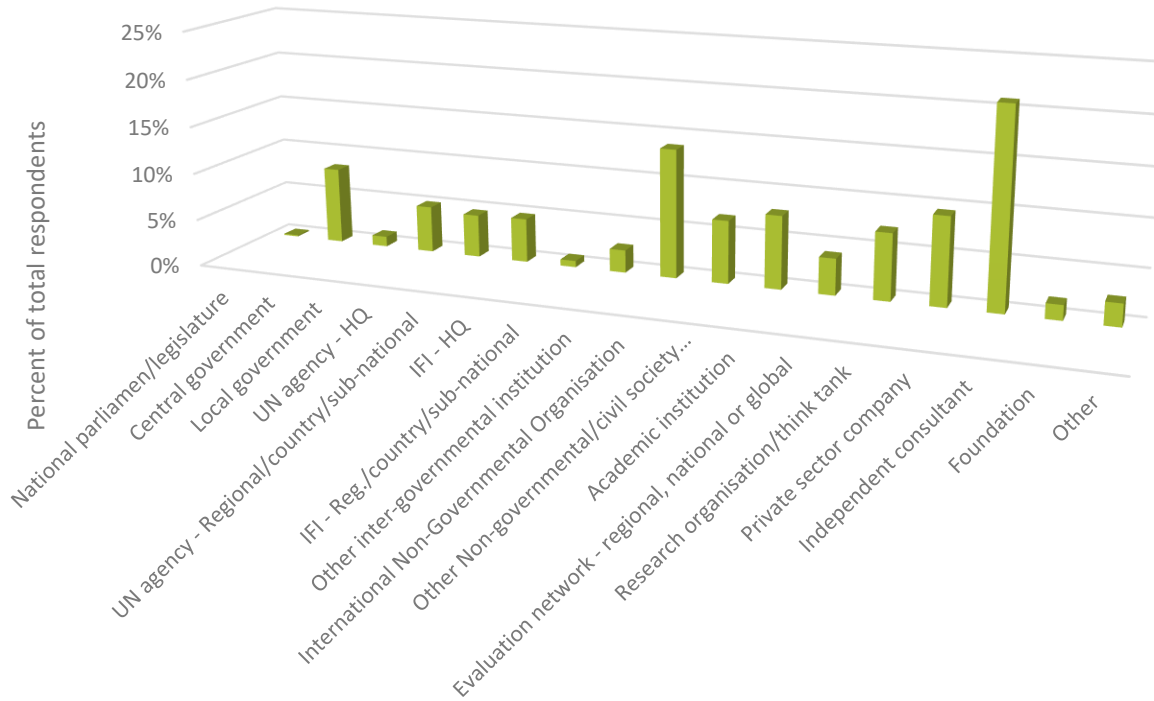
- **Independent consultants** – 21%

- **INGOs** – 14%

- **Private sector companies** – 9%

- **Central government/ academic institutions** – 8%

- **Research organisations/think tanks** – 7%

What role do respondents play in evaluations?

Pie chart values: 46, 35, 54, 67, 34, 42, 45, 24, 10

Legend:
- Designing Terms of Reference
- Contracting evaluation teams
- Managing evaluation processes
- Conducting evaluations
- Managing an evaluation function
- Presenting evaluations
- Evaluation research
- Study of evaluation
- Other
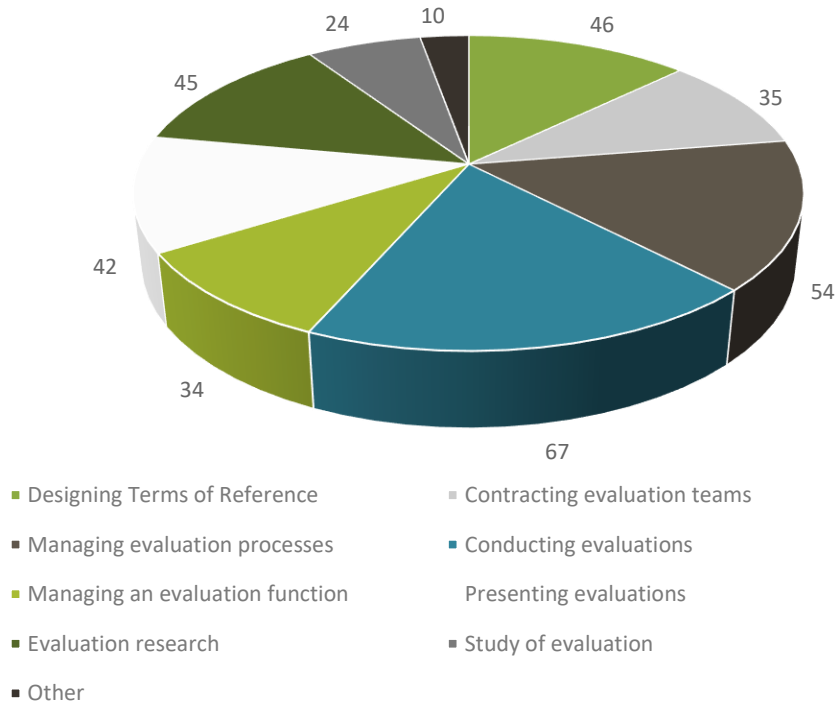
Respondents have **diverse/multiple roles in evaluations:**

- 67% **conduct evaluations**

- 54% **manage evaluation processes**

- 46% **design Terms of Reference**

- 45% **conduct evaluation research**

- 42% **present evaluations** to senior management/Executive Boards

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

Perceived Strengths
Universal acceptance & use across evaluation profession & beyond – creation of a common language & understanding; a normative framework
Standardisation/consistency – provide scope for comparability/synthesis across evaluands/contexts
Comprehensiveness/completeness – coverage of key areas required for accountability and learning
Simplicity and clarity – readily understandable, clear formulation
Neutrality - acceptable across cultures/political contexts
Universality - applicable to different evaluands; institution types; policy/intervention areas etc
Utility – provide relevant information to support improvement/change
Results focus – emphasise the importance of results at different levels
Adaptability/flexibility - can be tailored for different evaluands, contexts etc
Conciseness/feasibility – 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 believes, 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 2: How well are the current five criteria implemented?

560 total respondents

Implementation of evaluation criteria
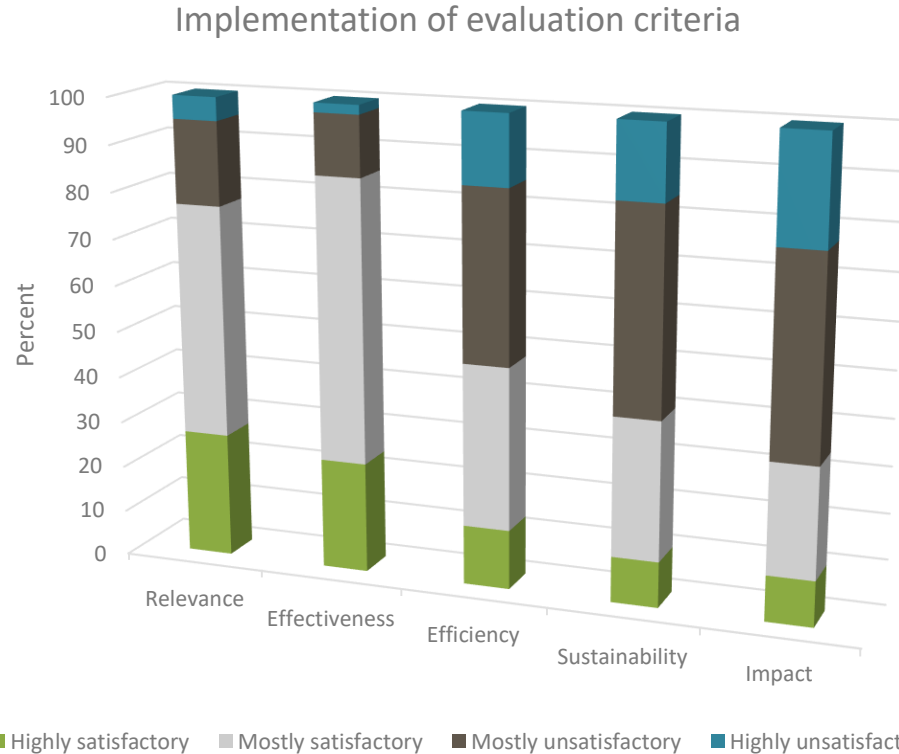
Variable perceived implementation

- *Most satisfactory implementation* (over 75% of respondents) **–** **Relevance/Effectiveness**

- *Less satisfactory implementation* – (over 40% of respondents) - **Efficiency, Sustainability**

- *Least satisfactory' implementation* (34% of respondents) **– Impact.**

Views 2: Current criteria implementation

Overall weaknesses in implementation *(most frequent responses)*:
- Applied **uncontextualised for the evaluand**
- Applied **mechanistically,** without **tailoring/adaptation for the evaluand** (a 'straitjacket')
- **All five applied**, whether appropriate or not
- Applied **in isolation,** rather than relationally
- Evaluator **capacity gaps** hinder implementation

Illustrative quotes:

'The criteria are useful, the problem resides in the way the criteria have been used.'
'One cannot blame the tools when they are misused.'
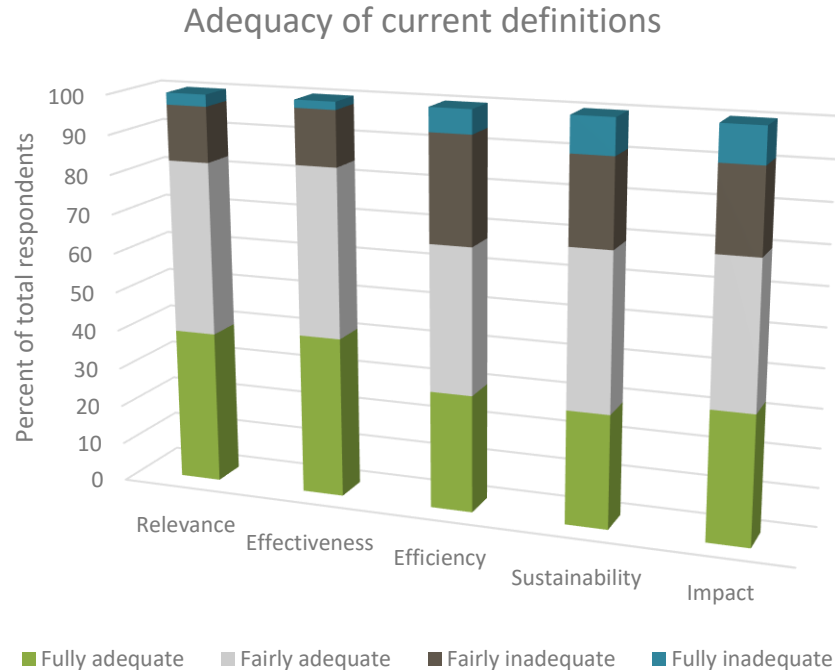*'The main problem with the criteria is when people treat them as a checklist…. If they are approached in humility as a preliminary guide about what constitutes worth and merit, then they work fine.'*

Individual criteria *(most frequent responses)*	
Relevance	• Shallowly assessed as 'alignment' with policies/strategies • Insufficient consideration of relevance of design to needs & subsequent adaptation
Efficiency	• Methodological weaknesses in implementation • Data gaps/limitations • Contradictions with Glossary definition
Effectiveness	• Appropriateness/quality of targets not assessed
Impact	• Criterion confused with specific methodologies • Data gaps/limitations
Sustainability	• Types of sustainability undefined • Vulnerable to 'evaluator opinion' alone • Assessed solely as 'funding continuity'

Views 3: How adequate are the current definitions of the criteria?

Answered: 510

Adequacy of current definitions

Legend: Fully adequate · Fairly adequate · Fairly inadequate · Fully inadequate

- **Majority of respondents** (over 65% for all five criteria) consider definitions for all five criteria 'Fully/Fairly' adequate

- **Relevance and Effectiveness:** Over 80% consider both definitions 'Fully/Fairly adequate'

- **Efficiency,** 67% consider definition 'Fully/Fairly adequate', 33% 'Fully/Fairly Inadequate'

- **Sustainability and Impact -** Over 69% consider both definitions 'Fully/Fairly adequate'; 9% consider definitions 'Fully inadequate'

Views 3: Current definitions of the criteria

Overall views on current definitions *(most frequent responses):*
The definitions are generally adequate but would benefit from refinement. Specifically:

- Re-align for closer alignment with the SDGs
- Explicitly recognise complexity
- Encompass interconnectedness
- Ensure adaptable for context/evaluand/evaluation type
- Provide greater specification where appropriate

Illustrative quotes

- ''*I actually think the definitions are still pretty useful. And I'm apprehensive of changing them too often or as fashions/SDGs come and go.*'
- '*Mostly adequate but need agreement on underlying guidelines.*'
- '*The definitions need adapting. They are too narrow and linear.*'

Comments on individual definitions *(most frequent responses)*

Relevance
- Encompass specification of 'relevant to whom/what'
- Consider all context dimensions – political/environmental/governance/social/cultural etc
- Embed equity considerations
- Include quality of design incl. logic model/theory of change
- Integrate comparative advantage/partnerships
- Assess quality of policy/strategy to which evaluand aligned

Impact
- Embed the principle of 'contribution'
- Clarify that 'impact' is a result, not a methodology
- Differentiate between actual and prospective achievement of longer term results
- Differentiate between results at different levels
- Specify range of 'longer term results', to include poverty reduction, institutional change, behavioural change, environmental change, realisation of human rights etc

Views 3: Individual criteria definitions

Comments on individual definitions *(most frequent responses)*

Effectiveness
- Assess quality of intended results
- Include timeframe of results (short, medium, long-term)
- Include scope/scale of intended results/needs met
- Integrate results for inclusiveness/participation
- Include adaptive capacity/agility
- Include results for gender, equity, human rights

Sustainability
- Specify whether: <u>actual</u> or <u>prospective</u> sustained results
- Assess whether sustainability strategies were designed-in/implemented from the outset
- Include statement on whether results are potentially transformative
- Define sustainability for the evaluand (transition/recovery/transformative change)
- Specify dimensions of sustainability incl. financial/social/environmental/institutional/cultural
- Include replicability/scaleability

Efficiency
- Remove term 'economically' (unclear) from definition
- Ensure coherence with Glossary definition
- Specify application at different levels of results chain
- Clarify tools – cost:benefit, cost-effectiveness, cost-efficiency analysis
- Encompass broad understanding of 'value for money'

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

Answered: 466
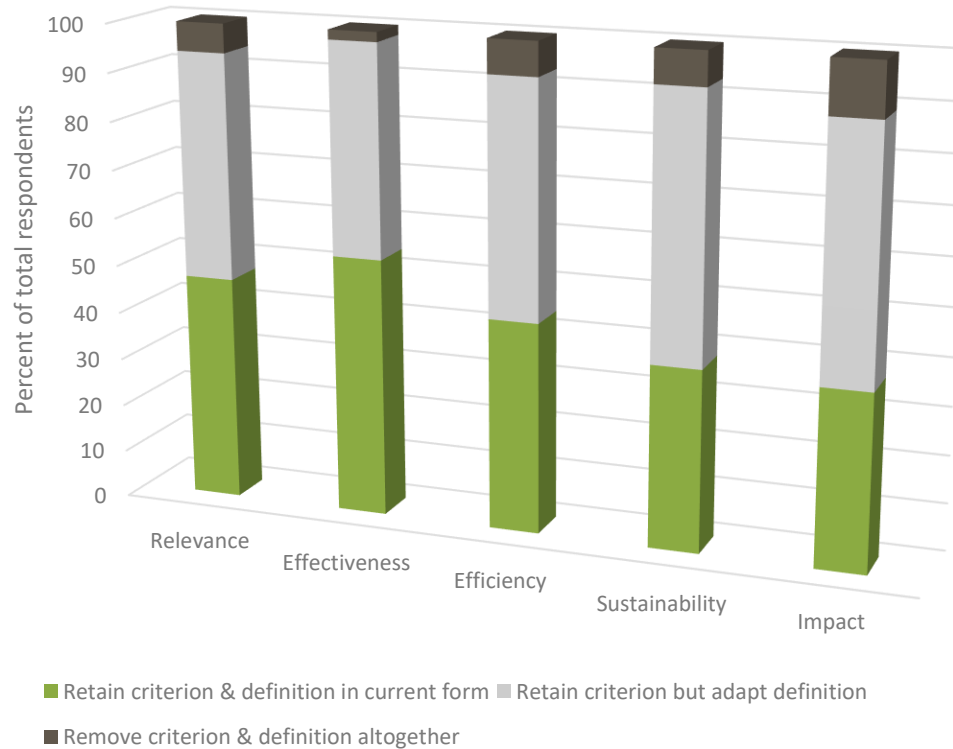
High majority of respondents preferred to *Retain criterion & definition in full* or *Retain the criterion but <u>adapt</u> definition* (over 89% of responses for all five criteria)

- **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 definitio*n (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 for*m (37%) or *Retain criterion but adapt definition* (52%). 11% preferred to *'Remove'* the criterion.

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' adaption
- 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 5: Potential role of humanitarian and peacebuilding criteria

Answered: 207
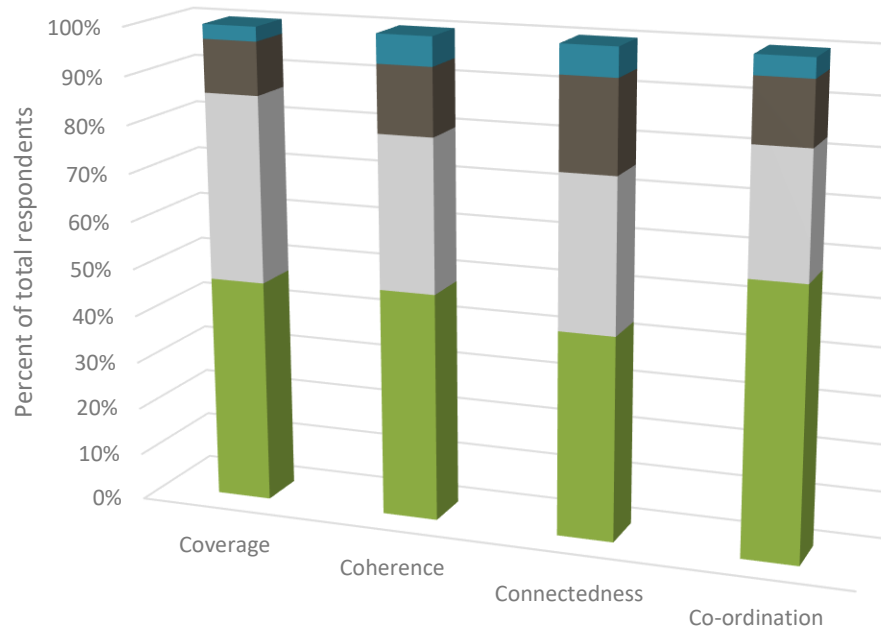
- **Majority of respondents:** Criteria 'very' or 'quite' useful/relevant' (over 75% for all four criteria)

- **Coverage and Co-ordination:** The most relevant/useful (*Coverage (86%) and Coordination (83%) very/quite useful/relevant):*

- **Coherence and Connectedness:** 80% and 75% respectively very /quite relevant/useful

- **Highest level of 'not at all useful/relevant'** only 6%, for Coherence and Connectedness.

Legend: Very useful/relevant ■ Quite useful/relevant ■ A little useful/relevant ■ Not at all useful/relevant

Views 5: Potential role of humanitarian and peacebuilding criteria

Respondents felt that humanitarian & peacebuilding criteria have potential relevance for development assistance – particularly as regards the SDG agenda and 'no one left behind'

Coverage - Linked to **Relevance/ Effectiveness** *(most frequent responses)*

Areas of relevance:
- Breadth of coverage (targeted population, geographic reach, access)
- Depth of coverage (impartiality, specific needs, inclusion).

Coherence – Linked to Relevance *(most frequent responses)*

Areas of relevance:
- *External coherence*
   - With policies/priorities of other actors including partner governments
   - With human rights, gender and equity considerations
- *Internal coherence*
   - Between activities, approaches, different elements of the intervention

Connectedness - Links to sustainability *(most frequent responses)*

Areas of relevance
- Links to recovery, transition, development activities

Co-ordination – Central to SDA agenda *(most frequent responses)*

Areas of relevance:
- Harmonisation, promotion of synergies
- Avoidance of gaps, duplication, resource conflicts.
- Co-ordination with national systems/institutional frameworks

Views 6: Proposals for additional criteria

- **28** additional criteria proposed
- …..But a common plea to retain a limited list

> 'Do not increase the number but integrate different aspects in them'
> 'Please do not add too many new criteria…as they will be standardised by many actors. We do not need 10 DAC criteria'
> 'Keep it simple!'

Proposed criterion	No, of suggestions *(where 10 or more)*
Equity, gender, human rights	57
Partnerships, synergies	49
Inclusiveness, stakeholder engagement	41
Co-ordination	26
Environment/climate change	25
Design rigour	22
Advocacy	17
Innovation	17

Proposed criterion	No. of suggestions *(where 10 or more)*
Agility/adaptive management	17
Coherence	15
Ownership	15
Internal coherence	14
Scaleability, replicability	14
Risk-willingness	11
Learning	12

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.'