CGIAR | Independent Advisory and Evaluation Service

Conducting and Using Evaluability Assessments in CGIAR

CGIAR EVALUATION GUIDELINES 2022

Citation

Independent Advisory and Evaluation Service (IAES). (2022). CGIAR Evaluation Guidelines. Conducting and Using Evaluability Assessments within CGIAR. Rome: Independent Advisory and Evaluation Service.

Design and layout

Khania Curtis Design and Creative Direction

Name of Guidelines:

Guidelines for Conducting and Using Evaluability Assessments within CGIAR

Purpose:

To operationalize the 'Evaluability' and related standards in the CGIAR Evaluation Framework, with the objectives to improve evaluability, ensure cost-effectiveness and contribute toward a continuous learning culture in CGIAR.

Audience:

The primary audience is evaluators, evaluation managers and commissioners of an evaluability assessment in CGIAR. Secondary audience- those responsible for providing inputs into evaluations in the CGIAR system; and stakeholders to the CGIAR's evaluation practice. Users in other agricultural research-for-development (AR4D) contexts may find this document useful too.

Framework and Policy Reference:

This guidance supports the CGIAR Evaluation Framework and the CGIAR Evaluation Policy (2022) and should be read in conjunction with other evaluation-related guidelines.

Acknowledgments

This guidance note has been co-developed by the Evaluation function of the CGIAR Independent Advisory and Evaluation Services (IAES): Svetlana Negroustoueva, Evaluation Function Lead, and Keith Child, consultant, under the overall direction of Allison Grove Smith, Director of the IAES. This guidance note draws on the practical experiences of evaluability assessments conducted by Rick Davies, Keith Child, and Sara Vaca. The authors would like to acknowledge the CGIAR MEL community of practice, specifically Helen Ashtul, Amanda Wyatt, Enrico Bonaiuti, Alessandra Furtado, and Julien Colomer whose feedback enriched the guidance. The earlier drafts benefitted from steering by Julia Compton (International Water Management Institute - a CGIAR Research center) and Roberto La Rovere (member of the IAES/Evaluation ERG (Evaluation Reference Group). The broader IAES team in Rome helped ensure smooth editorial and design support, technical assistance and administrative processes related to the development of this guidance.

It is encouraged that feedback and learning from the roll-out and application of this Evaluability Assessment (EA) guidance is shared with IAES/Evaluation (IAES-Evaluation@cgiar.org) as custodian of the document. The next version will build on direct experiences and evolving industry standards on EAs to best ensure effective and cost-efficient evaluations in CGIAR and similar contexts.

Contents

Boxes

Figures

Tables

Abbreviations

AR4D	Agricultural Research for Development
CRP	CGIAR Research Program (prior to 2022)
EA	Evaluability Assessment
EF	Evaluation Framework
EP	Evaluation Policy
ERG	Evaluation Reference Group
ISDC	Independent Science for Development Council
MELIA	Monitoring, Evaluation, and Learning and Impact Assessment
MELCOP	Monitoring, Evaluation, and Learning Community of Practice
MR	Management Response
PRMF	Performance and Results Management Framework
PRMS	Performance and Results Management
QoS	Quality of Science
ToC	Theory of Change
ToR	Terms of Reference

1 Introduction

In 2022, the CGIAR System Council and Board approved a CGIAR Evaluation Framework and a revised Evaluation Policy. Their development followed a consultative, inclusive, and iterative approach that included strategic and operational discussions with, among others, CGIAR governance and management, the CGIAR Monitoring, Evaluation, and Learning Community of Practice (MELCOP), and an external peer review. The Evaluation Framework and Policy define evaluation as the systematic and objective assessment of an ongoing or completed project, program, initiative or policy, or operational modality in CGIAR, its design, implementation, and results. Rigorous, independent, external evaluations are foundational to CGIAR's effort to inform the design of interventions, provide actionable evidence to support management and governance decisions, and ensure a high level of accountability to donors. Meeting this potential, however, requires advanced planning and appropriate evaluative inputs. 'Evaluability' (Box 1) is one of 15 standards and principles of the Evaluation Framework – a reference point for professionalism within research-for-development evaluation that underpins how evaluation is conducted in CGIAR. (See Figure 2).

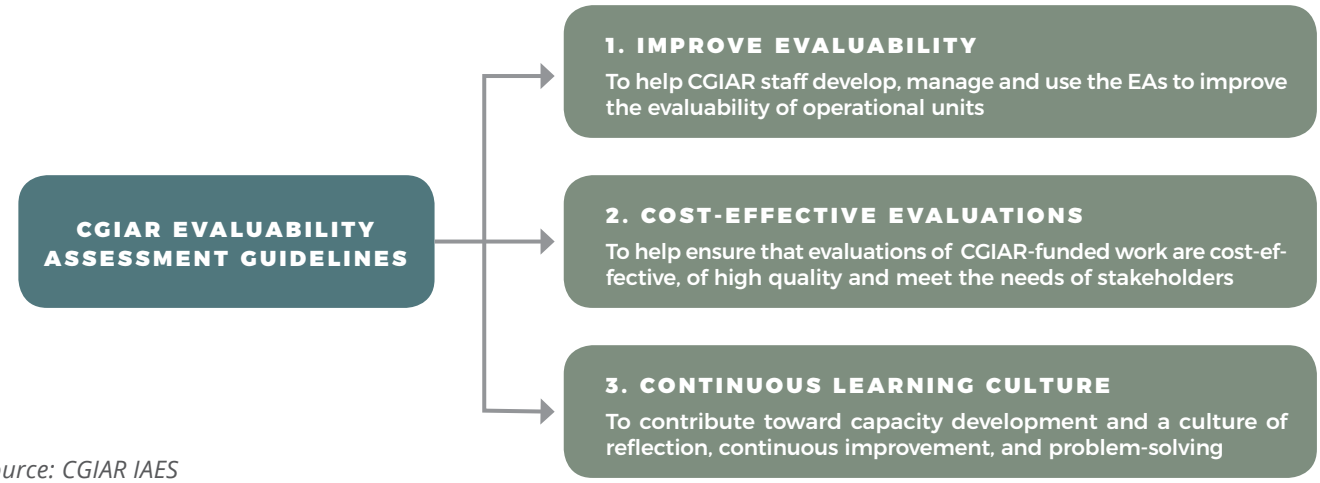
Box 1: *Evaluability Standard CGIAR Evaluation Framework, 2022*

> **Evaluability** refers to the extent to which an intervention can be evaluated in a reliable and credible fashion; the concept is central to a culture of results. A strong focus on **evaluability** at the design stage facilitates overall 'measurability', monitoring, and subsequent evaluation.

1.1 Purpose

These guidelines intend to facilitate better evaluation outcomes by ensuring that structures, processes, and resources for credible, rigorous, and useful evaluations are in place through the use of evaluability assessments. Alignment to evaluation standards and the application of evaluation criteria (Appendix One) underpin the theory of change (ToC) for evaluation practice in CGIAR (Evaluation Framework 2022, Section 2). Assessing evaluability enables the strategic consideration of all the elements of an evaluation before one is conducted so that evaluations are more likely to serve as a powerful management, learning, and accountability mechanism. Accordingly, these guidelines have three main objectives: improve evaluability, ensure cost-effectiveness and contribute toward a continuous learning culture (Figure 1).

Figure 1: *Purpose of the Evaluability Assessment guidelines*



Source: CGIAR IAES

1.2 Rationale

The CGIAR Evaluation Framework and the Evaluation Policy are supporting CGIAR transformation as stated in the 2030 CGIAR Research and Innovation Strategy and the Performance and Results Management Framework (PRMF). CGIAR has set out Initiatives that are prioritized areas of investment and research that seek to significantly contribute toward meeting the Sustainable Development Goals. The Initiatives include Regional Integrated Initiatives to provide pooled solutions to locally identified needs. CGIAR has articulated five Impact Area Platforms to foster critical thinking and the use of evidence to improve research impact. Each unit and layer of organization creates new challenges and opportunities for co-design, co-learning, and co-creation of synergies greater than any one initiative can deliver. Effective forward-looking processes and performance evaluations based on CGIAR's nested theories of change (ToC)[1] are needed to deliver results effectively and improve both the quality and usefulness of evaluations prioritized as lessons learned.

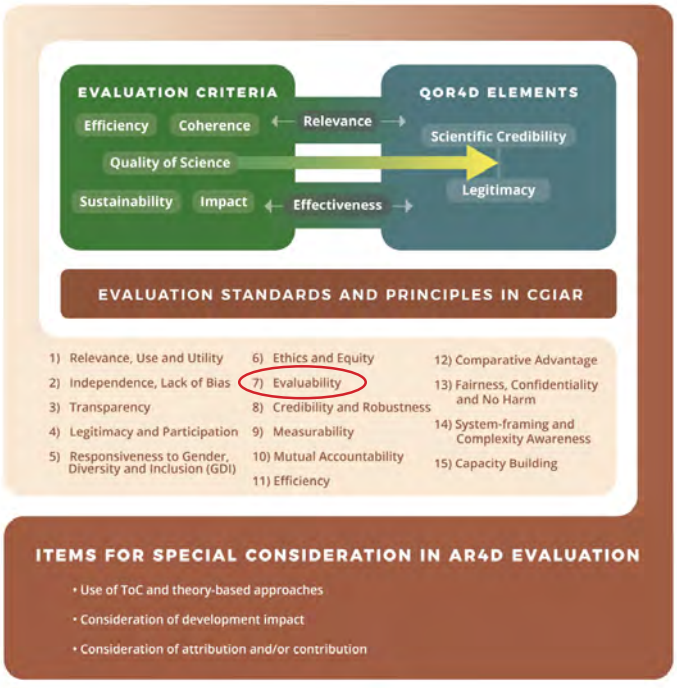
The Synthesis of Learning from a Decade of CGIAR Research Programs (2021)[2] pointed to implications for the evaluability of CGIAR programming. The 2021 synthesis revealed that the underlying ToCs of the Research Programs were in many cases not well articulated or comprehensive "with varying quality" and not clearly linked to the results chain. While there was verifiable evidence with regards to outputs, the synthesis findings also highlighted evidence and data gaps associated with measuring higher-level outcomes (aka System-Level Outcomes, and other development outcomes in the Strategic Research Frameworks) and weaknesses in the monitoring, evaluation, learning, and impact assessment (MELIA) systems. These impeded the comparability of the evidence from the 43 CGIAR evaluations and evaluative reviews. To strengthen MELIA metrics and systems tailored to CGIAR's mission, the synthesis recommended concrete steps to improve MELIA systems and evaluation practice[3].

The Independent Science for Development Council (ISDC) in its external review of the initiative proposals raised similar concerns. The ISDC review notes that "... integration of metrics into MEL plans needs to be more explicit (pg 7)[4]." Furthermore, the individual initiative proposal reviews conducted by the evaluation function in IAES often echoed 2021 synthesis recommendations and further highlighted the need for an explicit focus on quality MEL plans and coherence of results frameworks.

These findings challenge CGIAR to ensure that MEL systems are fit-for-purpose and that results are verifiable towards accountability and facilitate learning. Evaluations provide actionable evidence for management and governance decisions, facilitate learning and ensure accountability to funders and other stakeholder groups.

CGIAR's Evaluation Framework (EF) and Evaluation Policy (EP) (2022) articulate standards and criteria for process and performance evaluations to support CGIAR to deliver its mission and implement its 2030 Strategy against commitments in the CGIAR's Performance and Results Management Framework 2022-2030 (Figure 2) Towards operationalizing "Evaluability" standard, an evaluability assessment would help identify potential problems and solutions to allow CGIAR interventions sufficient time to adjust before an evaluation begins.

Figure 2: *CGIAR evaluation standards, principles and evaluation criteria with QoR4D elements*



Source: CGIAR Evaluation Framework 2022

[1] CGIAR uses a 'nested' ToC approach that progresses from a very high level of abstraction at the overall portfolio level, with increasing detail to Action Areas, Initiatives and Platforms, to Work Packages and finally to individual projects (PRMF 2022-2033, 7).

[2] The Synthesis covers two phases of CGIAR Research Programs: 2011-16 and 2017-21.

[3] For example: strengthened MELIA metrics and nested ToCs with clear impact pathways (Rec. 27); increased use of mixed methods designs in evaluations that expressly acknowledge the contributions of other research and innovations actors (Rec. 29); Improved MELIA coverage of cross-cutting themes (Rec. 30); expanded technical assistance in MELIA for research managers, scientists, and partners (Rec. 31).

[4] The reviews noted, among other things, that many of the initiative ToCs lacked causal linkages, thus making uncertain that the proposed work packages were sufficient to produce the desired End of Initiative outcomes. Additionally, many of the initiatives lacked quantifiable metrics required to measure progress, thus, jeopardizing monitoring, learning and evaluation plans. https://iaes.cgiar.org/sites/default/files/pdf/ISDC-12-Proposal-Review-_0.pdf

2 What is an Evaluability Assessment?

A high-quality and timely evaluability assessment is the best way to ensure that eventual evaluations generate actionable recommendations on topics of most importance to internal and external stakeholders. During an evaluability assessment, judgments are not made about the intervention and what has been achieved, but about the possibility of making such judgments and their likely utility. In doing so, it helps clarify ex-ante the feasibility, scope, and value for money of an evaluation. The duration and costs of an evaluability assessment are therefore considerably lower than those of an evaluation of the same initiative.

The CGIAR definition of evaluability assessments (Box 1) builds on that of the Development Assistance Committee of the Organization for Economic Co-operation and Development (OECD DAC, 2002).

Evaluability assessments and evaluations have two things in common: they both assess the quality of an intervention's design (relevance) and its internal consistency and contextualization (coherence). While they may share some points of analytical focus, it is essential to remember that they are separate but related processes. As such, an evaluability assessment is also a pre-evaluation tool, acting as a health check on MELIA components related to performance and impact – the most significant variables in determining evaluation readiness.

Box 1: *CGIAR definition of Evaluability Assessment*

An Evaluability Assessment is an early review of a proposed activity in order to ascertain whether its objectives are adequately defined, results verifiable, and evaluation questions answerable. An Evaluability Assessment establishes whether and how an intervention can be evaluated reliably and credibly.

An Evaluation is the systematic and objective assessment of an ongoing or completed project, program, initiative or policy, its design, implementation, and results. (CGIAR 2021).

3 Why Conduct an Evaluability Assessment?

An evaluability assessment is the first step to ensuring that an evaluation will be cost-effective, result in a useful learning process, and in findings that inform decisions and improve performance. It helps clarify the feasibility, scope, and value for money of an evaluation (Box 2). Even under the best of circumstances, evaluation readiness should never be assumed.

Experience suggests many reasons why an evaluation may fail to meet expectations including, but not limited to:

- Poor timing that may result in an evaluation being conducted either too early or too late to provide actionable recommendations to inform strategic planning or course correction, make a policy change, or inform a follow-on proposal for new funding.
- Performance data that may be available but of low quality, rendering evaluative judgments open to contestation.
- A lack of demand or weak buy-in from stakeholders may mean that the evaluation is regarded as irrelevant to their needs or a burden and is, consequently, underutilized.
- The lack of a skilled Evaluation Manager and other MEL staff who are required to facilitate the evaluation roll out and follow-up.
- An evaluation budget that may not be sufficient to answer the key questions for the stakeholders, thereby minimizing the use of evaluation findings and recommendations in decision-making.

Box 2: *Twelve key reasons to conduct an evaluability assessment*

Conducting an evaluability assessment will help to:

1. Quality assure adequacy of MEL-related during the intervention design process.
2. Strengthen and adjust the theory of change (e.g., to explicitly address QoR4D, to focus on how the intervention will lead to expected outcomes in a particular context, to stress test how results will be measured along the impact pathways).
3. Inform the design or revision of the results framework.
4. Enable consensus building among evaluation stakeholders to manage expectations of what an evaluation will produce.
5. Inform evaluation design, i.e., scoping and Terms of Reference.
6. Solidify lines of enquiry and evaluation criteria, including evaluation questions and to identify questions that may not be answerable.
7. Facilitate early engagement and buy-in of stakeholders to the evaluative process.
8. Facilitate access to and availability of core documentation for evaluation.
9. Ensure alignment to Evaluation Framework's standards.
10. Guide integration of specific topics unique to the evaluation.
11. Ensure timelines and efficiency of the evaluation to prevent wasting resources.
12. Provide management evidence to advocate for necessary MEL resourcing and capacity.

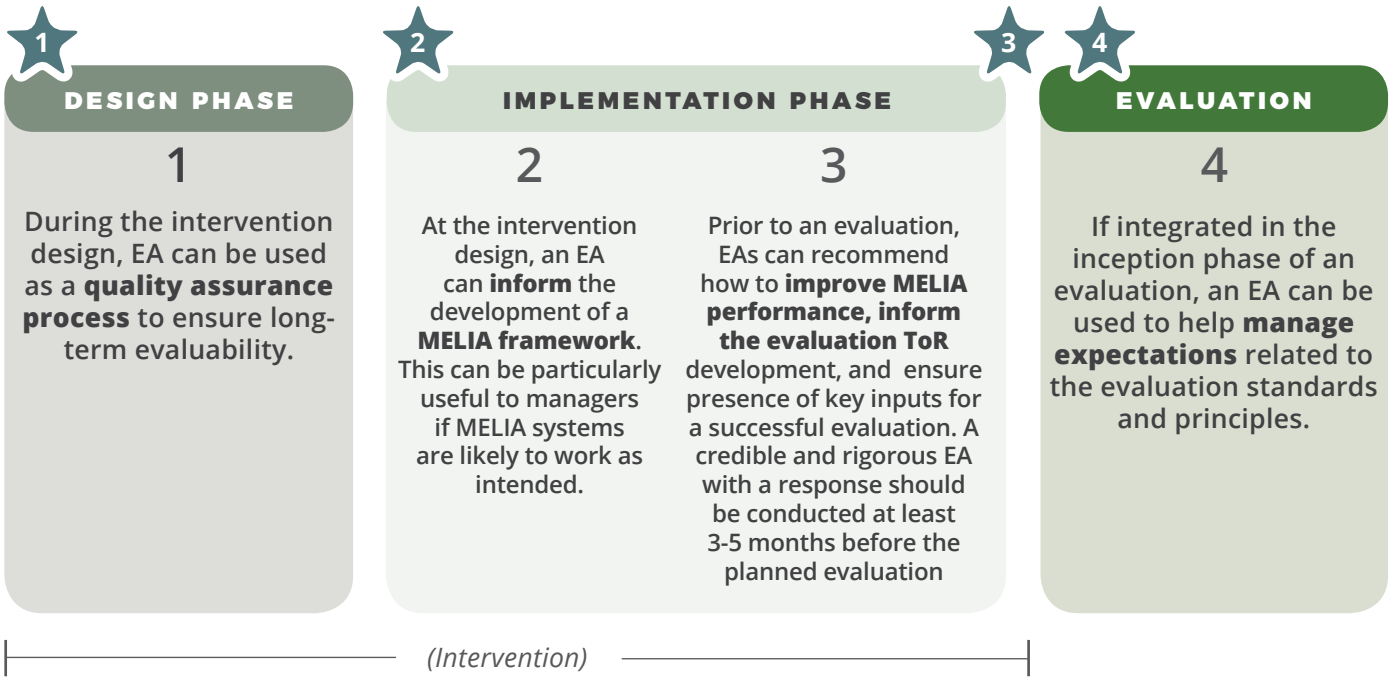
An evaluability assessment is a systematic and reasonable approach to mitigate these and other potential obstacles.

4 What to Assess for Evaluability and When?

Across the CGIAR portfolio, evaluability assessments are key for producing evaluations aligned with the standards and practices of the Evaluation Framework. Consistent with the scope of the Evaluation Policy, EAs can be conducted on any object of evaluation such as initiatives, Platforms, thematic areas of CGIAR research and corporate entity, or any activity area (e.g., an Initiative Work Package or individual project).

Prioritization among a substantial number of interventions, partner options, and engagement strategies, should point to which interventions are most crucial to conduct an EA. Within CGIAR, the need for an evaluability assessment would depend on whether evaluations are independent or largely independent (see Table 1).

Figure 3: *Phases of the project cycle by use of an evaluability assessment*



4 What to Assess for Evaluability and When?

Independent evaluations: Commissioned by CGIAR System Council through the IAES (as per the approved multi-year evaluation plan[5]), independent evaluations assume a stand-alone evaluability assessment. Based on the Evaluability Assessment results, the IAES may advise that a comprehensive evaluation will need to be postponed or canceled.

Largely independent (or decentralized evaluations[6]): Evaluations of Initiatives and Platforms or other thematic evaluations commissioned by management. An example could be the evaluation of the CGIAR Research Program on Policy, Institutes, and Markets (PIM)[7]. The EA for this type of evaluation can be conducted at any time during the funding cycle (phases 1-4 – see below); stand-alone EAs are not required but encouraged for largely independent and/or decentralized evaluations.

The timing may also align to four phases in the funding cycle (Figure 3) when it is most useful to conduct an assessment: three as a stand-alone assessment and the fourth as an integrated component of an evaluation inception report.[8]

For either type of evaluation, independent or largely independent, managers should weigh their decision before conducting an evaluability assessment. While an EA is useful for many reasons (Box 3), there are justifiable reasons for not conducting one, including but not limited to:

- The scale of the intervention is so small that the EA cost would be out of proportion to its benefits.
- The intervention is known to be unevaluable in its current form and remedies are ongoing. In this case, effort should be directed toward ensuring that remedies are fully implemented.
- From the intervention design and start-up stage (Phases 1-2), managers and funders have strong evidence and concur that the intervention is evaluable; thus making the EA before commissioning the evaluation unwarranted.

Table 1: *Expectations for conducting evaluability assessment classified by the commissioning entity in CGIAR*

	Independent, external evaluations	Largely independent/ decentralized evaluations
Object/scope	Initiatives and Platforms, or other thematic and corporate level evaluations	Initiatives and Platforms, or other thematic evaluations
Commissioner	System Council	Management
Through IAES/ Evaluation function	Yes	No
Included in multi-year evaluation plan	Yes	No
EA expected	Yes	Left to the discretion of a commissioning entity
Management response	Yes	Left to the discretion of a commissioning entity

Source: CGIAR IAES

[5] CGIAR Advisory Services: 2022-24 Workplan at the time of development of this guidance.

[6] 'Largely independent' evaluations, also known as 'decentralized', are commissioned by management.

[7] An example is a 2021 PIM Partnership Evaluation, commissioned and conducted by IFPRI/PIM CRP https://ebrary.ifpri.org/utils/getfile/collection/p15738coll2/id/135071/filename/135282.pdf . See also the related blog.

[8] While accepted as an option in the evaluation industry, the integration of an evaluability assessment into an evaluation inception phase limits its use for enhancing MEL resourcing, structures and processes. There is also a potential conflict-of-interest in that the evaluability assessor/ evaluation team may not recommend postponing an evaluation or identify 'limitations' that cannot be mitigated.

5 Who Should Commission and/or Conduct an Evaluability Assessment?

The consideration of staffing requirements for an evaluability assessment is driven by who commissions it and its timing.

An evaluability assessment can be conducted by CGIAR staff or by an external consultant with either approach being desirable for different reasons. On the one hand, a deep working knowledge of CGIAR can foster efficiency while on the other hand, an external consultant may produce a more independent, rigorous, and objective assessment, but with less efficiency. In the case of EAs commissioned internally by CGIAR management, designated MEL staff in CGIAR may implement it with or without external consultant support.

For independent, external evaluations[9] commissioned by CGIAR System Council, an evaluability assessment would be a stand-alone activity conducted in advance of an evaluation. For these EAs, the IAES will retain independent external consultants to conduct the assessment. For largely independent/decentralized evaluations, the independence of the assessor for an evaluability assessment is less important than for an evaluation as the assessor renders no evaluative judgment, thereby making it feasible to assign CGIAR staff to conduct the EA. Internal CGIAR staff can be used for a condensed EA commissioned by management or as part of scoping or inception in the independently commissioned evaluation (detailed further in Section 6, Table 3).

The level of effort required to conduct an EA can vary significantly. Table 2 provides a very rough estimate of time allocation and leadership roles using as an example two types of EA. In practice, the level of effort and division of labor required to conduct an EA are context-specific and will vary from case to case. However, regardless of the context and type of evaluation and commissioner, the assessment will require honest reflection, dedicated time, and effort from CGIAR staff (see Roles and Responsibilities – Evaluation Policy). It should be considered that while an EA requires time and money, in most cases, these are expenses that would otherwise be invested in an actual evaluation.

Table 2: *Level of effort and roles by two types of EA Frameworks*

Source: CGIAR IAES

	Condensed EA Framework (Table 3)	Full EA Framework (Appendix 3)
EA Domains/questions	5 domains/14 questions	6 domains/40 questions
Estimated duration, assuming full time level of effort	1-2 weeks, excluding travel	5-6 weeks, excluding travel
Leadership and Phase in project cycle (Figure 3) example	Largely independent/decentralized evaluations commissioned by management and conducted by CGIAR staff (Phase 1-3)	Independent, external evaluation commissioned by IEAS/Evaluation and conducted by an Assessor and (Phases 3-4)
Step 1: Identify Purpose & Scope	Management of evaluand with MEL staff; Potential Quality Assurance by IAES/Evaluation	IAES / Evaluation
Step 2: Identify & involve stakeholders	MEL staff of evaluand	MEL staff facilitate stakeholder involvement / Assessor
Step 3: Apply selected EA Framework	MEL staff of evaluand	Assessor applies Framework
Step 4: Collect evidence	MEL staff of evaluand	MEL staff facilitate document collection; Assessor reviews documents
Step 5: Assess readiness; make recommendations	MEL staff of evaluand	Assessor judges evaluation readiness
Step 6: Use EA results	Management with MEL staff; Potential Quality Assurance by IAES/Evaluation	Management with MEL staff; CGIAR Governance

[9] Coordinated by the Evaluation Function within the Independent Advisory and Evaluation Service (IAES).

5 Who Should Commission and/or Conduct an Evaluability Assessment?

In terms of desired qualifications to conduct an EA, a mix of evaluation and subject matter expertise is recommended, with prior evaluation experience in the CGIAR context and overall familiarity with CGIAR being highly desirable. Evaluation expertise is necessary to address methodological issues around data and its analysis while subject matter expertise is needed to assess the plausibility of the expected effects of interventions, the quality of evidence, and the potential usefulness of findings. In short, internal, desk-based assessments, this mix of expertise may not be possible, but in longer independent, sometimes field-based assessments, involving stakeholder consultations, mixed expertise is recommended[10].

Like an evaluation, an evaluability assessment requires a **Terms of Reference** (ToR) which would capture elements presented in Box 4. The ToR facilitates the conduct of the assessment by setting out recommended steps to follow.

Box 4: *Core elements of the evaluability assessment Terms of Reference*

1. **Context and background on evaluand**
2. **Purpose**
3. **Scope**
4. **EA questions following the Evaluability Framework**
5. **Design and approach**
6. **Work plan**
7. **Staffing requirements**
8. **Data and documentation**

See Appendix Two for a detailed example.

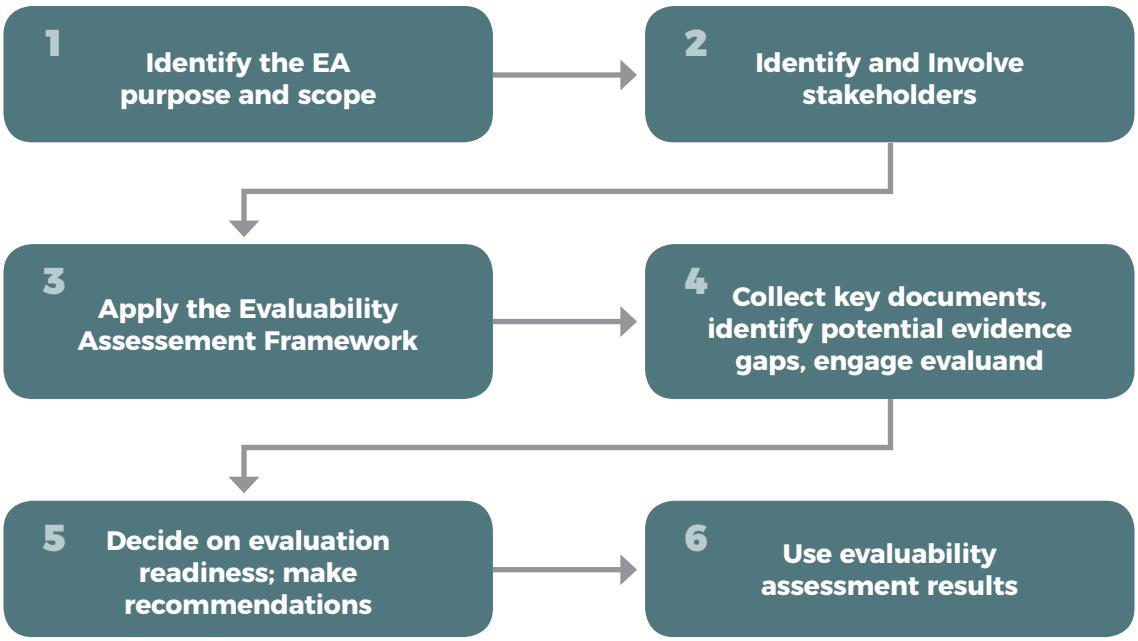
[10] Evaluability Assessments in Austrian Development Cooperation (entwicklung.at)

6 How to Conduct an Evaluability Assessment in CGIAR?

6.1 A Six-step Process

These guidelines set out a six-step approach (Figure 4) to conducting an evaluability assessment. Assessors, potentially in consultation with a commissioner, should exercise judgment at each step to align the EA with contextual realities and the needs and expectations of management.

Figure 4: *Six-step evaluability assessment (EA) process*



Step One: Identify the Purpose and Scope of the Evaluability Assessment

The first step is to identify the evaluability assessment's purpose and scope. To a large extent, this will depend on when it is conducted in terms of the four phases of the funding cycle (Figure 3). For assessments conducted in anticipation of an evaluation (phases 3 and 4), an evaluability assessment would help narrow the focus of the evaluation to a subset of the seven evaluation criteria (Figure 2) and prioritize questions by key interests of stakeholders. Appendix Two provides a list of topics that should be included in the Terms of Reference (ToR) for an EA.

Because CGIAR interventions are complex in the AR4D context, the following questions are ideally resolved at this step:

Purpose: Is the EA intended to be a precursor to an evaluation or a quality check on the evaluand? See Box 3 to guide and narrow down the list of key reasons to conduct an evaluability assessment, towards formulating its purpose.

Scope: Is the evaluation assessment limited to a specific geographical focus, and or time frame (e.g., security threats may prevent an evaluability assessment around election periods; data availability might not be possible for the entire intervention)?

If the answer to these questions is unclear, an EA can help to resolve them.

Step Two: Identify and Involve Stakeholders to Make Evaluability Assessments Both a Product and a Learning Process

CGIAR management, scientists, funders, and research partners have a vested interest in and responsibility for an evaluation as do scaling partners, country governments, beneficiaries, and many others. As with evaluations, evaluability assessments can create largely unnecessary concerns among the stakeholders of an intervention being assessed. To be credible and legitimate, transparency is critical.

Consistent with standards and principles set out in the Evaluation Framework (including Relevance, use, and utility; Transparency; and Legitimacy and Participation), conducting an evaluability assessment begins by identifying the needs of potential evaluation participants, what they want to know and how they will use evaluation results. Additionally, elaborating the preferred format for communicating results to diverse stakeholder groups begins here. To manage expectations and pave the way for the assessment, the scope, purpose, and expected outputs should be communicated to stakeholders in advance.

Including stakeholders fosters engagement and buy-in to the eventual evaluation process itself so it is important to involve the right people from the start through stakeholder mapping to:

- Identify all stakeholders.
- Classify and map their expectations and roles in the evaluation process: establish who should participate, their interest in the evaluation results, and their role in the decision-making process.
- After the evaluability assessment is completed, it is advisable to conduct webinars and workshops to share results with stakeholders, particularly when stakeholder groups are numerous.

Step Three: Apply the Evaluability Assessment Framework

The CGIAR Evaluability Assessment Framework (Appendix Three) includes the key domains and criteria for conducting an assessment, the primary point of which is to answer questions about the evaluability of the intervention under examination along the **Evaluability Assessment Framework Domains** (Figure 5). The goal of the assessment is **not to make an evaluative judgment** but merely **to assess adequacy of the theory of change, establish the interests of the stakeholders, and whether available and potentially available data would make it possible to formulate a judgment** given the initiative timeframe and resources. An assessment will help determine which evaluation type is warranted, the status of implementation, and the availability of evidence used to make evaluative judgments. External factors that may be obstacles to implementation and performance monitoring-related data collection should also be identified, such as a late start-up, security issues or natural disasters, and political or economic instability. The hurdles to implementation need to be understood and accounted for from the very start of an evaluation.

The Evaluability Assessment Framework is generally applicable to the entire CGIAR portfolio of investments although some objects of evaluation may require it to be adapted to their specific needs (e.g., an evaluation of a thematic area of research that spans multiple Initiatives may not have an 'intervention logic'. It is likely that a conceptualization of expected outcomes and differences in these would be available and may prompt reconstructing the theory of change to guide an actual evaluation).

The level of effort for this step is undoubtedly the most labor intensive and adequate resources (e.g., time, expertise, and budget) need to be in place to review available evidence. This step aims to compare the 'implementation reality' to the 'design of the object of evaluation.' A wide variety of data collection methods can be employed, including document review, focus group interviews, indicator analysis, observation, and so on. For example: For Domain B (MEL systems and resources), an early check on MEL resourcing, and a review of selected performance data will help identify if there are sufficient data toward selected findings for drawing evaluative conclusions regarding the achievement of targets.

For largely independent evaluations, a condensed Evaluability Assessment Framework can be applied (Table 3). The condensed Framework includes only six out of seven domains (14 core assessment criteria as opposed to 40 questions) found in the Evaluability Framework available in Appendix Three. The condensed Framework can be used for largely independent, decentralized, or other evaluations commissioned by management – it can be employed more rapidly and with fewer resources (see Table 1). For independent evaluations, the full Evaluability Framework will be used.

Figure 5: *Evaluability framework domains*

Table 3: *Core evaluability domains and assessments – condensed framework*

Core Evaluability Domains	Core Assessment Criteria summarized	Y/N Comment
A. Intervention logic: To be evaluable, an intervention must clearly describe what it hopes to achieve and how. Intervention logic represents the overall logical integrity of the intervention and should be supported by a robust ToC and a body of evidence that lends credibility/plausibility to the ToC.	**1. Theory of change:** (a) Is there an explicit ToC (or logical framework) that describes the intervention's expected results and impact pathways? (b) Are the ToC model and narrative well aligned? (e.g. the narrative explains the model, and elaborates the causal logic with examples).	(a/b) Y/N
	2. Quality and quantity of evidence base: Is the ToC supported by a credible body of external evidence (primary or secondary)?	Y/N
	3. Clarity of intervention additionality, comparative advantage[11] and spheres of control: Has evidence been provided around specific claims been made about the intervention's contribution, and that of other actors, to the achievement of the desired results? (Spheres of control and influence of an intervention)	Y/N
	4. Feasibility: Are the ToCs' causal logic realistic and feasible to achieve within the timeframe and resources allocated?	Y/N
	5. Complexity: Are there complex relationships between different intervention components that will make attribution of results difficult to assess?	Y/N
B. MEL systems and resources: To be evaluable, an intervention must have a credible plan to track its contribution to outcomes. It should reflect a vision of how monitoring and evaluation activities will fulfill accountability, delivery, and learning needs. The MEL system must generate relevant and quality data, most often by an intervention's indicators. A baseline is a necessary starting point against which to assess intervention performance and results.	**6. Quality of MEL Framework:** Does the intervention have a MEL framework in place that is fit to generate evidence to support all the key events in the ToC and generate the data in support of the Results Framework and PRMF?	Y/N
	7. Information resources and system in support of MEL: Are there designated MEL personnel at the level of the evaluand?	Y/N
	8. Quality of indicators or other measures: Does the intervention results framework include indicators that are appropriate to evidence its ToC from outputs to impacts?	Y/N
	9. Quality of Baseline: Does the intervention have baseline evidence against targets towards its objectives?	Y/N
C. Gender, diversity, and inclusion: The CGIAR is committed to the inclusion of women, youth, and socially excluded and vulnerable groups.	**10. Clarity of partners and end-user groups:** Are partners and other stakeholders clearly defined, within spheres of control and/or influence, along with how their interests may coincide or conflict?	Y/N
	11. Data disaggregation (gender, youth, other): Do existing data allow for data disaggregation according to targeted cross-cutting groups?	Y/N

[11] Towards the development of the related concept note, at the time of finalizing the EA guidelines the following blog was considered https://iaes.cgiar.org/isdc/news/effectively-using-concept-comparative-advantage-within-cgiar

Table 3 cont'd

Core Evaluability Domains	Core Assessment Criteria summarized	Y/N Comment
E. Context and environment: To be evaluable, an intervention must be accessible to evaluators.	**12. Accessibility:** Is there anything about the timing of a planned evaluation that would make it difficult/ impossible to conduct (e.g., seasonality, budget allocations, public holidays, local elections)?	Y/N
F. Management and key stakeholder engagement and support: To be evaluable, management and stakeholders must be active participants in the evaluation process.	**13. Return to management and other key stakeholders:** Do evaluation criteria and potential questions address the issues of importance to stakeholders?	Y/N
	14. Demand from and Participation of Key Stakeholders: Do stakeholders understand their role and potential contribution to an evaluation?	Y/N

Note: Evaluability Domain D 'Long-term evaluability' is highly contextualized to the project cycle and timing of evaluability assessment, and it is therefore not included as a core domain.
Source: CGIAR IAES

Timing effects the relative attention to:
1. ToC (early in design/intervention)
2. M&E systems (midway)
3. Stakeholder questions (prior to evaluation)

Step Four: Collect Key Documents and Identify Potential Evidence Gaps, re-engage Evaluand as necessary

While activities carried out in Step Three focus on assessing whether available and potentially available data make it possible to carry out an evaluation, Step Four prioritizes the availability of evidence. In most cases, the evidence also represents the initial documentation for evaluation scoping and design. As part of this step, and in alignment with Step Two, it is important to re-emphasize the role of stakeholders in providing timely and complete evidence. CGIAR interventions are complex, involving large numbers of people and activities, captured in a variety of evidence sources[12]. Collecting core and other evidence and information is a time-consuming and arduous activity for the commissioning office and evaluation team- tasked with actual evaluation. Making the relevant documentation available beforehand, facilitated by the results of the evaluability assessment, would save time and money and allay possible misunderstandings regarding the scope and objectives of the evaluation.

A crucial step of an evaluability assessment process and its outputs is to highlight the importance of facilitating access to and availability of relevant evidence for the eventual evaluation and identify relevant documentation and data sources to identify potential gaps. For efficiency purposes, only the core documents need to be considered and compiled at this stage, along with the names and contact information of key stakeholders. The use of an online document archive (SharePoint or E-Library) is a convenient way to manage this process while considering that:

A CGIAR Performance Results Management System (PRMS) or related modules (i.e. Results Dashboard or other CGIAR dashboards) should contain key performance-related information.[13] However, outdated or incomplete data are not uncommon, so it is good to review key documents (Box 5) during the evaluability assessment (Domain B: MEL systems and resources) to ensure they are up-to-date and complete within the planned time scope of the evaluation.

If the use of 'Quality of Science' evaluation criterion is likely within the planned objectives of the evaluation[14], special consideration should be given to information and data for assessing outputs, consistent with the data assets indicator of the PRMF. Costs of commissioning an independent bibliometric and/or altimetric analysis need to be embedded in recommendations for evaluation design and timing[15].

[12] In alignment with the SIMEC-endorsed CGIAR Technical Reporting Arrangement (2022); it describes the content, timing, format, standards and scope of technical reporting applicable to all CGIAR Initiatives; and Figure 9 illustrates documentation to be considered for stage-gating.

[13] PRMS and its modules are under redesign during the elaboration of this document. The CGIAR dashboards and decision registers provide access to information on CGIAR results, financials, Trust Fund and governance decisions.

[14] Consult 'Evaluators' Guide: Applying the CGIAR Quality of Research for Development Framework to Process & Performance Evaluation"- under development parallel to this guidance.

[15] Science-Metrix & CGIAR Advisory Services Secretariat Evaluation Function (2022). Bibliometric Analysis to Evaluate Quality of Science in the Context of One CGIAR. Technical Note. Rome. in the Context of One CGIAR. Technical Note. Rome.
https://cas.cgiar.org/evaluation/publications/bibliometric-analysis-evaluate-quality-science-context-one-cgiar

Box 5: *Examples of key documentation*

- **Proposal** or strategy documents
- **Theory of change** (at all intervention levels, and including historic and current versions, if updated annually)
- **Results Framework** or other documents with articulated inputs, activities and outputs, desired outcomes and impacts and indicators
- **Plan of Work and Budget, Annual (and other) reports**
- **Project lists** (with related documentation)
- **Contact lists** for internal and external stakeholders and key informants
- Standard **indicator description sheets**
- **ISDC reviews**, funder review/comment on proposal as applicable
- Previous independent or other evaluations, studies, and impact assessments
- Access to key databases and data sources with relevant information, e.g. monitoring data and staff lists, access to Web of Science for bibliometrics, dashboards
- Internal audit or other assurance-related documents
- Governance-related information (Independent advisory bodies, structure, ToRs, composition, meeting minutes)

*Items in **bold** indicate the minimum, core required list, while other documents listed should be collected if available.*

Step Five: Decide on Evaluation Readiness and Make Associated Recommendations

The Evaluability Assessment Framework (Appendix Three) is used to determine if an intervention can be evaluated and its results determine whether and how an evaluation should be conducted.

At the onset, and in the process of applying the EA framework, questions would be asked about how the findings about each of the domains aggregate and fit with each other, as findings from one area of analysis are likely to have implications on others. In practice, the process is likely to be much more cyclical and iterative. Stakeholders are likely to provide information about relevant documents, relevant documents can provide information about relevant stakeholders, both will provide information about information systems, and vice versa.

A summary judgment derives from aggregating the assessment of the individual questions and judgments by domain. The use of a checklist with a narrative explanation is suggested to facilitate the triangulation of results. The checklist should be considered as a guide rather than a determinant to move forward with an evaluation.

At this stage, sufficient information has been collected to decide on evaluation readiness. If the EA criteria have not been met, the results from an EA can be used to develop recommendations that will lead to a higher state of evaluation readiness in the future, meaning the evaluation can be conducted once appropriate mitigation measures to potential limitations are in place. An evaluability assessment may also help refine evaluation questions or the evaluation criteria to only those that can reasonably be answered and applied rather than cancel an evaluation outright (Table 4).

Table 4: *Decisions support framework using stoplight system, full EA Framework*

	Indicator	Decision	Follow-on Steps/Recommendations
Evaluability assessment criteria met	**Yes**, to all 14 core questions	Proceed with evaluation	Feed-in lessons and evidence from EA in programming &/or evaluation design: develop evaluation ToR, adapt ToC and measurements, and convene stakeholders in a decision-making committee, such as an Evaluation Reference Group.
Evaluability assessment criteria partially met	**Yes**, to all 14 core questions + Y to at least half of the other core questions even if with caveats	Postpone the evaluation	Address issues raised, reassess & refine key evaluation design parameters (criteria/questions, timing, resources)
None of the evaluability assessment criteria are met	**No**, to all 14 core questions + N to more than half of the other questions	Cancel the evaluation	The intervention is not ready for an evaluation until deficiencies are remedied, e.g., institutional bottlenecks that could undermine the evaluation, data quality that is insufficient to support meaningful evaluative judgments, or external conditions that make conducting an evaluation too difficult or dangerous. In the meantime, the recommendations can inform MEL-related adjustment and any other aspects aligned to the lines of inquiry as per the framework.

Source: CGIAR IAES

For a condensed framework, using core evaluability domains and assessments (Table 3) an alternate decision framework could be used with three or four different elements:

1. The type of decision – whether to proceed with an evaluation, and if so when. These possibilities are already described, but a bit more freedom could be given for example to indicate how long a postponed evaluation could feasibly take place.

2. The conditions that might apply to any of these decisions would need to be met before they could be implemented. These would refer to specific aspects of the evaluability, as listed in the EA framework.

3. The EA team's arguments for why these conditions need to be met.

4. Other improvements which could be made but which are not conditional.

Depending on the phase (Figure 3) and purpose of the evaluability assessment (Step 1) weighting of EA framework domains can be considered. In this case, a written justification should accompany the scoring matrix.

The CGIAR Evaluation Framework (2022) establishes the overall approach to process and performance evaluations. An evaluability assessment may also clarify what approaches and types of evaluation to conduct (Box 6). The MELCOP Glossary provides CGIAR-appropriate definitions, while alternative resources, such as the BetterEvaluation platform, should be consulted for detail on various types and approaches[16]. While it may be premature to assign an evaluation approach and methodology aligned to the standards and principles, it may be possible to give some indication of which methods might be considered in an evaluation design and later an evaluation inception report.

[16] In the TORs for many evaluations, typically, a wide range of evaluation questions represent the diversity of different stakeholders interests in the evaluation. One particular method would rarely address these different questions. Mixed methods approach is always preferred, with some methods being more appropriate to some questions. If an EA is inclined to make recommendations about particular methods, such recommendations would need to be done with caution, encouraging creative combination of methods with different strengths and weaknesses rather than any single method "solutions '

Box 6: *Examples of potential evaluation types and approaches*

Formative evaluation: Formative evaluation ensures that a program or program activity is feasible, appropriate, and acceptable before it is fully implemented. It is usually conducted when a new program or activity is being developed or when an existing one is being adapted or modified. Formative evaluation is generally any evaluation that takes place before or during implementation with the aim of improving the design and performance of an intervention.

Developmental evaluation: An evaluation approach that can assist social innovators develop social change initiatives in complex or uncertain environments. Development evaluation originators liken their approach to the role of research and development in the private sector product development process because it facilitates real time, or close to real-time, feedback to program staff thus facilitating a continuous development loop.

Democratic Evaluation: Various ways of doing evaluation in ways that support democratic decision making, accountability and/or capacity.

Outcome Mapping: An impact evaluation approach which unpacks an initiative's theory of change, provides a framework to collect data on immediate, basic changes that lead to longer, more transformative change, and allows for the plausible assessment of the initiative's contribution to results via 'boundary partners'.

Step Six: Using Results of an Evaluability Assessment

How an evaluability assessment is used depends on who commissions it and when it is commissioned during the funding cycle of an intervention (Figure 3). In line with the stakeholder mapping exercise in Step 2, recommendations are directed and communicated to pre-identified users (e.g., funders, or the management team of the object of evaluation which is most likely to be an Initiative or Platform).

6.2 Document a Management Response/Action Plan

Intervention managers should engage with the evaluability assessment results, particularly if the assessment makes recommendations and produces actionable advice that must be remedied before an evaluation occurs. Where recommendations are made, corrective steps must be initiated as soon as possible.

Consistent with the CGIAR Evaluation Policy, a Management Response to evaluations underpins evaluation quality, rigor, and credibility[17]. By responding appropriately to recommendations, managers reaffirm their commitment to performance-based management and set the stage for a useful evaluation experience. A **Management Response is required** for all evaluability assessments for evaluations commissioned by CGIAR System Council through the independent evaluation function. **This ensures the efficiency of eventual evaluations that will be conducted, by engaging intervention managers to assure any remediations around the 6 domains have taken place**. In CGIAR context, a management response to other EAs (e.g., those supporting a largely independent evaluation commissioned by management) are encouraged but not required.

[17] A co-designed guideline supporting Management Engagement and Response establishes a clear road map to operationalize sections 6 and 7 of the Evaluation Policy (2022). Additional resource is a process note on 'Developing, tracking and reporting on Management Responses to Technical Evaluations' by CGIAR management (under development at the time of finalizing this guidance) with accompanying online tool-Management Response Tracker.

6.3 Inform Evaluation Terms of Reference

Results of the evaluative assessment and the Management Response, if available, can and should inform the development of a Terms of Reference for an evaluation[18] including:

1. **Context:** providing information on differences in the expectations and interests of stakeholders in the intervention. Institutional, cultural, and physical constraints that might affect the progress of an evaluation should also be mentioned here.

2. **Theory of change:** enabling the Terms of Reference to highlight the strengths and weaknesses of the Theory of Change.

3. **Evaluation criteria and questions:** informing the choice of which evaluation questions are included and excluded and referencing any consensus and differences of opinion about the relative priority of the different evaluation questions.

4. **Data sources:** point out strengths and weaknesses in the availability and quality of potentially relevant data, as well as the capacities and limitations of the monitoring and evaluation system(s) in generating this information. This information should be made available within the Terms of Reference for an evaluation.

5. **Evaluation approach and methods:** information provided in the above three sections will have implications for the kinds of evaluation approaches and methods that can be used. The Terms of Reference should highlight the most important of these implications. This will include visible opportunities and what seems unlikely to be feasible at the time of the assessment. The evaluation team should be left with room to make choices within these bounds since they will be responsible for the actual implementation and delivery of results.

6. **Evaluation Limitations:** Evaluability assessment results, potentially combined with learning from the scoping exercise, would enable identification of constraints envisaged for the evaluation and limitations of the suggested approach and methodology, such as resource limitations, openly discussed together with their implications for evaluation. The Inception report would provide detail on how the team plans to over come and mitigate them. Notably, when applicable a Management Response to EA would need to be referenced, especially when potential limitations cannot be mitigated.

[17] Guidance on Evaluation Terms of Reference is available from the evaluation function under the Independent Advisory and Evaluation Service. It would be adapted for specific needs of an evaluation.

7 Potential Risks

As an evaluative activity, an evaluability assessment needs to be managed to avoid some commonly known risks, including but not limited to:

Stakeholder confusion: The difference between an evaluative assessment and an evaluation may not be apparent to all stakeholders, resulting in a lack of cooperation or a sense of insecurity. To avoid confusion, communication about the purpose of the evaluative assessment is critical. Effective communication with stakeholders is essential 'groundwork' before an assessment begins.

Evaluation overload: CGIAR managers and staff may feel that the extra step of an evaluative assessment is too time-consuming, expensive, or intrusive. Again, effective communication is vital. CGIAR is an institution that actively seeks to promote a culture of evaluation. Managers can help mitigate the risk of evaluation overload by explaining the purpose of the evaluability assessment and of evaluation to their staff as an essential component for maintaining CGIAR's reputation for AR4D excellence.

Perceived lack of efficiency: An evaluability assessment represents an additional cost to an evaluation although usually, it can be conducted for a small fraction of the overall evaluation budget. Nevertheless, any additional costs may be resented during a time of fiscal belt-tightening. To help mitigate this risk, it is essential to emphasize that an evaluability assessment can save money in the long run, particularly if it prevents a poorly conceived or executed evaluation.

Delay: When an EA is conducted in anticipation of an evaluation, it may raise concerns that need to be addressed before an evaluation can be conducted. In rare cases, this could lead to a delay in the planned timing of the evaluation. If an evaluation needs to be rescheduled, it is important that management thoroughly document their response to the recommendations of the evaluability assessment. There may be no easy way to lessen the impact of a delayed evaluation except to emphasize the value of addressing potential problems before an expensive and formal evaluation is conducted.

Table 5: . *Evaluability assessments: potential risks and mitigation strategies*

Risks	Mitigation Measure
Stakeholder confusion: The difference between an EA and an evaluation may not be apparent to all stakeholders, resulting in a lack of cooperation or a sense of insecurity.	Stakeholder engagement and effective communication.
Efficiency: An EA represents an additional cost to an evaluation.	Value for Money: EA ensures cost-savings during an actual evaluation, costs for evaluations will fall accordingly.
Evaluation overload: CGIAR managers and staff may feel that the extra step of an EA is too time-consuming.	Stakeholder engagement and effective communication.
Delay: When an EA is conducted prior to an evaluation, it may raise concerns that need to be addressed before an evaluation can be conducted.	Adaptive management.

Appendix 1: CGIAR Evaluation Criteria (Evaluation Policy)

Relevance: The extent to which the intervention's objectives and design respond to the needs, policies, and priorities of users/clients and global, regional, and country partners/institutions and continue to do so if circumstances change. Consistent with the QoR4D framework, attention is given to the importance, significance, and usefulness of the work implemented in the problem context, associated with CGIAR's capacity to address the problems.

Effectiveness: The extent to which the intervention is achieved, and/or is expected to achieve, its objectives, and its results, including any differential results across subgroups of users/clients. Consistent with the QoR4D framework and in the CGIAR context, this criterion considers the extent to which research is positioned for use and has generated knowledge, products, and services with high potential to address a problem and contribute to innovations, outcomes, and impacts. Effectiveness, therefore, implies that research has been designed, implemented, and positioned for use within a dynamic theory of change, with appropriate leadership, capacity development, diversity of research skills, and support to the enabling environment to translate knowledge into use and to help generate desired outcomes.

Coherence: The compatibility of the intervention with other interventions in a country or a sector or within CGIAR; its overall fit. Internal coherence addresses the synergies and interlinkages between the intervention and other interventions carried out within CGIAR, and the consistency of the intervention with the relevant international norms and standards to which CGIAR adheres. External coherence considers the consistency of the intervention with other actors' interventions in the same context—that is, its complementarity, harmonization, and coordination with others, its value-added, and its avoidance of duplication of effort.

Efficiency: The extent to which the intervention delivers, or is likely to deliver, results in an economical and timely way—that is, the overall use of resources. "Economical" refers to the conversion of inputs (funds, expertise, natural resources, time, etc.) into outputs, outcomes, and impacts in the most cost-effective way possible compared with feasible alternatives in the context. "Timely" delivery is within the intended timeframe or a timeframe reasonably adjusted to the demands of the evolving context. This criterion may include assessing operational efficiency (how well the intervention was managed).

Quality of Science: The QoS evaluative criterion pertains to scientific credibility and legitimacy. The definition of the criterion derives from the QoR4D frame of reference, which records CGIAR's System-wide agreement on the nature and assessment of research quality. The QoR4D describes research quality according to four key elements: relevance, scientific credibility, legitimacy, and effectiveness.[19] Relevance and Effectiveness are treated as separate evaluation criteria above.

Sustainability: The extent to which the net benefits of the intervention continue or are likely to continue. This criterion focuses on the continuation of benefits, not on external funding, and highlights the multidimensional nature of sustainability.

Impact: The extent to which the intervention has generated or is expected to contribute to generating significant positive or negative, intended or unintended higher-level effects. Impact addresses the ultimate significance and potentially transformative effects of the intervention.

[19] A co-designed guideline on evaluating the quality of science in CGIAR details the approach and methods for operationalizing the Quality of Science evaluation criterion of this Policy.

Appendix 2: Terms of Reference for Evaluability Assessments

In addition to specific information required by the contracting party (e.g., payment modalities, submission guidance, timeline, etc.), an evaluability assessment ToR should also include the following:

1. Context
   a. The economic, social, and political context in which the intervention is being evaluated.
   b. Important background information about the intervention to be evaluated (e.g., changes in strategic focus, stage in the funding cycle, etc.).

2. Purpose
   a. The reason an evaluability assessment is being conducted.
   b. The primary audience for the evaluability assessment.

3. Scope
   a. Define the object of the evaluability assessment in terms of geography, timeframe, funding, or implementation partners.

4. Evaluability Assessment questions
   a. The Evaluability Framework outlined in this guidance document (Appendix Three)
   b. Additional evaluability questions that are specific to the intervention.

5. Design and approach
   a. The type of information sources that will be used for the assessment.
   b. Expectations around stakeholder consultation.
   c. Risk and potential mitigation measures.
   d. Types of outputs that are expected (e.g., inception report, final report, summary presentation of findings, list of recommendations, etc.)

6. Work plan
   a. A description of the key stages of the evaluability assessment.
   b. The total number of working days, subdivided by deliverables.
   c. Start and stop dates, including a timeline for deliverables.

7. Staffing requirements
   a. Who is eligible to conduct the evaluability assessment (e.g., internal or external assessors)?
   b. Qualifications and experience.
   c. Conflict of interest and eligibility constraints.

8. Data and documentation
   a. Description of data sources.
   b. Key contact persons.
   c. Reference to this and other guidance documents published by the CGIAR evaluation function under Advisory and Evaluation Services.

9. Annexes
   a. Any additional documentation relevant to the evaluability assessment, including URLs to online documents.

Appendix 3: Evaluability Framework

The CGIAR Evaluability Framework (Table 6) includes the key domains and criteria for conducting an Evaluability Assessment. The Framework is generally applicable to the entire CGIAR portfolio of investments although some objects of evaluation may require that the Framework be adapted to their specific needs (e.g., an evaluation of a thematic area of research that spans multiple Initiatives may not have an 'intervention logic').

The Evaluability Framework employs a simple stoplight scoring system (red = indicates a serious problem; yellow = a minor problem that can be remedied; green = fit for the Evaluation). The scoring system is for quick reference. The stoplight system should not be used on its own but is intended to accompany a narrative answer to the questions raised under each domain in the Framework. The narrative in the body of the evaluability assessment should provide deeper insight into each domain, sufficient to warrant its score based on comprehensive answers with evidence, where applicable.

Depending on the phase (Figure 3) and purpose of the evaluability assessment (Step 1) weighting of EA framework domains can be considered.

Default core criteria with higher relative importance (and weight) are reflected in the condensed framework-underlined assessment criteria (as per Table 3). These five domains bear disproportionately on the evaluability and, therefore, on the overall decision support framework using stoplight system. The absence of any of the core sub-criteria identified for a given domain will result in an "unsatisfactory/unmet" status for the overall criterion. As an example, for assessment criterion A – Intervention Logic, under 'Theory of Change', two core assessment sub--criteria have been identified: (a) and (b). If any of these core sub-criteria is assessed as "No" (i.e. sub-criterion not met), the overall rating of criterion "Theory of Change" will be "unsatisfactory", regardless of the assessment given to the other sub-criteria (c and d).

If weights are assigned, a written justification should accompany the scoring matrix.

Table: *Evaluability Framework*

Domain	Assessment Criteria	Indicator / Core Crtierion	Stoplight Scoring System
A. Intervention logic: To be evaluable, an intervention must clearly describe what it hopes to achieve and how. Intervention logic represents the overall logical integrity of the intervention. This logic should be supported by a robust ToC and a body of evidence that lends credibility/ plausibility to the ToC.	**Theory of change:** a. Is there an explicit ToC (or logical Framework) that describes the intervention's expected results and impact pathways?[20] b. Are the ToC model and narrative well aligned? (e.g. the narrative explains the model, and elaborates the causal logic with examples). c. Have ToC assumptions that are essential to the working of the ToC been explicitly stated, in sufficient detail they can be assessed? d. Are there sufficient causal linkages to plausibly suggest that intervention activities are sufficient to produce desired outcomes (e.g., End of Initiative outcomes)?	1	
	Quality and quantity of evidence base: e. Is the ToC supported by a credible body of evidence (primary or secondary)?	2	
	Clarity of intervention additionality, comparative advantage[21] and spheres of control: f. Have specific claims been made about the intervention's contribution, and that of other actors, to the achievement of the desired results (Spheres of control and influence of an intervention) which could be assessed?	3	
	Feasibility: g. Are the ToCs' causal logic realistic and feasible to achieve within the timeframe and resources allocated?	4	

[20] Underlined assessment criteria highlight only those core criteria that are reflected in the condensed form of the framework.
[21] Towards the development of the related concept note, at the time of finalizing the EA guidelines the following blog was considered https://iaes.cgiar.org/isdc/news/effectively-using-concept-comparative-advantage-within-cgiar

Appendix 3: Evaluability Framework cont'd

Domain	Assessment Criteria	Indicator / Core Crtierion	Stoplight Scoring System
A. Intervention logic: ***cont'd***	**Complexity:** h. Are there complex relationships between different intervention components that will make attribution of results difficult to assess? i. Is the ToC appropriately aligned and linked with other nested ToCs to realize broader strategic goals (e.g. organization, country)? j. Is there an explicit written acknowledgment of how complex change processes are expected to occur (i.e. nonlinearity, emergence, adaptation, feedback loops, etc.)?	5	
B. MEL systems and resources: To be evaluable, an intervention must have a credible plan in place to track its contribution to outcomes. The plan should include a unified vision of how M&E activities will fulfill accountability, delivery, and learning needs. The MEL system must generate relevant and quality data. Most often, this data is defined by an intervention's indicators. Having appropriate indicators that are aligned with desired results is essential. A baseline is a necessary starting point against which to assess intervention performance and results.	**Quality of MEL framework:** a. Does the intervention have a MEL framework in place that is fit to generate evidence to support all the key events in the ToC and generate the data in support of the Results Framework and PRMF? b. Does the MEL framework specify a monitoring plan for indicators (specifying who, when and how indicator data will be collected)?	6	
	Information resources and systems in support of MEL: c. Are there designated MEL personnel at the level of evaluand? d. Is the MEL system generating the data in support of the Initiative/intervention Results Framework and PRMF (information system, excel)?	7	
	Quality of approach to learning: e. Are there mechanisms in place for making use of findings from MELIA products (reviews, evaluations, impact assessments) for decision-making? f. What evidence is there about quality (strategy, uptake) and use of learning opportunities in the past?	---	
	Quality of indicators or other measures: g. Does the intervention results framework include indicators that are appropriate to evidence its ToC from outputs to impacts? h. Do indicators include both human/social and environmental qualities? i. Are there indicators around science delivery and quality? j. Are the indicators SMART?[22]	8	
	Quality of baseline: k. Does the intervention have baseline evidence against its objectives? l. Is baseline data related explicitly to the intervention indicators?	9	
	Quality of results statements? m. Are results statements for the evaluand (output, outcome, et al., statements) sufficiently clear for success to be recognizable? n. Do the result "types" make sense? o. Are elements framed correctly? (e.g. outcome statements are actor-specific and presented in active language: i.e., answers the question 'who is doing what differently?')	---	

[22] S-specific, M-measurable, A-achievable, R-realistic, T-timebound.

Appendix 3: Evaluability Framework cont'd

Domain	Assessment Criteria	Indicator / Core Crtierion	Stoplight Scoring System
C. Gender, diversity, and inclusion: The CGIAR is committed to the inclusion of women, youth, and socially excluded and vulnerable groups. To be credible and legitimate, CGIAR research must be based on the inclusion of the end-users it hopes to reach.	**Clarity of partners and end-user groups:** a. Are partners and other stakeholders clearly defined, within spheres of control and/or influence, along with how their interests may coincide or conflict? b. Have important differences between end-user groups been identified, concerning differences in their expected roles and results?	10	
	Data disaggregation (gender, youth, other): c. Do existing data allow for data disaggregation according to targeted cross-cutting groups? d. For assessing inputs to QoS - is evidence in place about young and mid-career researchers?	11	
D. Long-term evaluability: Many of the impacts of CGIAR research will not be recognized until long after intervention delivery.	**Sustainability:** a. Have the expectations about the nature and duration of the sustainability of the intervention and/or its effects been made clear enough to be evaluable? b. Is there clarity on the linkages across the ToC towards the potential and actual sustainable development impact, as appropriate?	---	
E. Context and environment: To be evaluable, an intervention must be accessible to evaluators and key stakeholders in the evaluation.	**Accessibility:** a. Is there anything about the timing of a planned evaluation that would make it difficult/impossible to conduct (e.g., seasonality, budget allocations, public holidays, local elections)? b. Are there security or political issues that would make a planned evaluation difficult/impossible to conduct? c. Are there any geographical constraints on accessibility, either by the evaluation team or by local stakeholders in the intervention?	12	
F. Management and key stakeholder engagement and support: To be evaluable, management and stakeholders must be active participants in the evaluation process.	**Return to management and other key stakeholders:** a. Do evaluation criteria and potential questions address the issues of importance to stakeholders? b. Is there a commitment to learning from evaluation findings by any of the stakeholders? c. Do stakeholders have mechanisms and the capacity to learn from potentially negative evaluation findings? d. Is there likely to be a process for ensuring Management engagement and response to the evaluation findings?	13	
	Demand from and participation of key Stakeholders: e. Do stakeholders understand expectations about their role and potential contribution to an evaluation? f. Are stakeholders available and ready to participate in an evaluation?	14	

Appendix 4: Bibliography and Further Reading

Austrian Development Agency. 2022. Evaluability Assessments in Austrian Development Cooperation. Vienna, Austria. Austrian Development Agency. https://www.entwicklung.at/fileadmin/user_upload/Dokumente/Evaluierung/GL_for_Evaluability_Assessments.pdf

CGIAR. 2021. Monitoring, Evaluation, Learning and Impact Assessment (MELIA) Glossary. CGIAR Report No. 5. Rome, Italy. CGIAR https://drive.google.com/file/d/1v0O5wt4z3bgs_wCYa7H2FifTVSAXAVjl/view

CGIAR. 2022a. CGIAR Evaluation Policy. Rome, Italy. CGIAR. https://cas.cgiar.org/sites/default/files/pdf/CGIAR%20CAS%20Evaluation%20Policy_24.3.2022_v2.pdf.

CGIAR. 2022b. CGIAR Evaluation Framework. Rome, Italy. CGIAR. https://cas.cgiar.org/sites/default/files/pdf/CGIAR%20CAS%20Evaluation%20Framework_24.3.2022_rev%2014%20April%202022.pdf.

Davies R, Payne L. 2015. Evaluability Assessments: Reflections on a review of the literature. Evaluation. 21(2):216–231. United Kingdom. doi:10.1177/1356389015577465. http://dx.doi.org/10.1177/1356389015577465.

DFID. 2013. Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations. Working Paper 40. Cambridge, UK. Department for International Development https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf

MacPherson, R., Jersild, A., Cours, D., & Holo, C. 2022. Assessing the Evaluability of Adaption-Focused Interventions: Lessons from the Adaptation Fund. In (eds.) Uitto JI, Batra G. Uitto, J.I. (2022). Transformational Change for People and the Planet: Evaluating Environment and Development. Sustainable Development Goals Series. Springer, Cham. https://doi.org/10.1007/978-3-030-78853-7_1

OECD (2021), Applying Evaluation Criteria Thoughtfully, OECD Publishing, Paris, https://doi.org/10.1787/543e84ed-en

USAID. 2021. Evaluation Toolkit V2. Conducting an Evaluability Assessment for USAID Evaluations. Washington, DC. USAID https://usaidlearninglab.org/sites/default/files/resource/files/et-_eval_assess_final2021.pdf

Vaessen, J. 2017. Evaluability and why it is Important for Evaluators and Non-Evaluators. [Blog accessed 17 July 2022]. https://ieg.worldbankgroup.org/blog/evaluability

CGIAR | Independent Advisory and Evaluation Service

Contact

CGIAR Independent Advisory and Evaluation Service (IAES)
Via di San Domenico 1, 00153 Rome, Italy
Email: IAES@cgiar.org
URL: https://iaes.cgiar.org/evaluation

Follow the Independent Advisory and Evaluation Service on social media: