Assessing the Evaluability of Adaptation-Focused Interventions: Lessons from the Adaptation Fund

Ronnie MacPherson, Amy Jersild, Dennis Bours, and Caroline Holo

Abstract

Evaluability assessments (EAs) have differing definitions, focus on various aspects of evaluation, and have been implemented inconsistently in the last several decades. Climate change adaptation (CCA) programming presents particular challenges for evaluation given shifting baselines, variable time horizons, adaptation as a moving target, and uncertainty inherent to climate change and its extreme and varied effects. The Adaptation Fund Technical Evaluation Reference Group (AF-TERG) developed a framework to assess the extent to which the Fund's portfolio of projects has in place structures, processes, and resources capable of supporting credible and useful monitoring, evaluation, and learning (MEL). The framework was applied on the entire proj-

ect portfolio to determine the level of evaluability and make recommendations for improvement. This chapter explores the assessment's findings on designing programs and projects to help minimize the essential challenges in the field. It discusses how the process of EA can help identify opportunities for strengthening both evaluability and a project's MEL more broadly. A key conclusion was that the strength and quality of a project's *overall* approach to MEL is a major determinant of a project's evaluability. Although the framework was used retroactively, EAs could also be used prospectively as quality assurance tools at the pre-implementation stage.

Introduction

Background to Adaptation Fund

The Adaptation Fund was established by the Parties to the Kyoto Protocol (CMP) of the United Nations Framework Convention for Climate Change (UNFCCC) to finance concrete climate change adaptation (CCA) projects and programs in developing countries that are particularly vulnerable to the adverse effects of climate change. At the Katowice Climate Conference in December 2018, the Parties to the Paris Agreement (CMA) decided that the Adaptation Fund (AF) shall also serve the Paris Agreement.

R. MacPherson
Greenstate, Newtonmore, UK
e-mail: ronnie@greenstate.net

A. Jersild
Evaluation consultant and doctoral candidate in the
Interdisciplinary Program in Evaluation (IDPE) at
Western Michigan University (WMU), Kalamazoo,
MI, USA
e-mail: amy.c.jersild@wmich.edu

D. Bours (✉) · C. Holo
Technical Evaluation Reference Group of the
Adaptation Fund (AF-TERG), Washington, DC, USA
e-mail: dbours@adaptation-fund.org;
cholo1@adaptation-fund.org

Since 2010, the Fund has committed \$720 million in grants to more than 100 projects in developing countries, with projects working in a diversity of sectors including agriculture, disaster risk reduction, coastal management, food security, and urban development. The Fund provides grants to implementing entities that lead the development, implementation and monitoring of work on the ground, usually in partnership with other organizations. Implementing entities can be multilateral (e.g., UN and multilateral agencies), regional (regional development banks), or national (government ministries, national research institutions).

As with other comparable institutions, the Adaptation Fund uses evaluation as a tool for understanding project results, strengthening accountability, learning, and continuous improvement. An evaluation framework (AF, 2012) sets out the Fund's approach, defining objectives, requirements, roles, and processes that should be applied when evaluating Adaptation Fund supported projects. Central to this approach, implementing entities are required to commission independent final evaluations (AF, 2011a, b) for any Adaptation Fund-supported project, with independent midterm evaluations if a project is more than 4 years in length.

The Fund's evaluation function was initially outsourced. In 2019, the Fund internalized evaluation with the establishment of the Adaptation Fund Technical Evaluation Reference Group (AF-TERG). An early step by the AF-TERG was to commission a series of preliminary studies to inform and support the development of a multi-year work program. One of these studies was an evaluability assessment (EA) of the Adaptation Fund's portfolio of projects.

History and Purpose of Evaluability Assessment

Michael Scriven (1991) defines evaluability analogous to requiring serviceability in a new car. It may be thought of as "the first commandment in accountability," notes Scriven (p. 138). The technical use of the term originated with Joseph

Wholey (1979) and his colleagues at the Urban Institute in the 1970s as a response to the delays and low value found in summative evaluations of U.S. government programs. EAs were a means by which to examine a program's structure to determine whether it could lend itself to generating useful results from an outcome evaluation. They were also viewed as a preformative evaluation activity that was part of a cost-effective strategy in determining readiness for evaluation and enhancing use. For Wholey, EAs were the first of four tools in a "sequential purchase of information," including rapid feedback evaluation, performance monitoring, and impact evaluation (Wholey, 1979; Wholey et al., 2010, p. 82).

The Development Assistance Committee of the Organisation for Economic Co-operation and Development (OECD DAC, 2002) defined evaluability as "the extent to which an activity or a program can be evaluated in a reliable and credible fashion" (p. 21) with a focus on methods and design, in contrast to Wholey's stronger focus on utility from a cost perspective. The OECD DAC further described EAs as an "early review of a proposed activity in order to ascertain whether its objectives are adequately defined and its results verifiable" (p. 21). Scriven (1991) noted the possible confusion that EAs may pose with regard to taking the place of serious summative evaluation, or to support the greater tendency to rely on objectives-based evaluation (a pseudo-evaluative approach) when the time comes to evaluate (Stufflebeam & Coryn, 2014).

Wholey (1979) developed an eight-step approach to implementing EAs:

1. Define the program to be evaluated.
2. Collect information on the intended program.
3. Develop a program model.
4. Analyze the extent to which stakeholders have identified measurable goals, objectives, and activities.
5. Collect information on program reality.
6. Synthesize findings to determine the plausibility of program goals.
7. Identify options for evaluation and management.

8. Present conclusions and recommendations to management.

Over the years, Wholey's approach was modified and others have further elaborated and emphasized certain aspects while reducing the number of steps. Smith (1989), for instance, identified stakeholder awareness as a particularly vital part of EAs, noting the importance of perceptions about what a program is to accomplish and whatever defined needs there may be for evaluative information on a program, whereas Rutman (1980) focused on methods and the feasibility of achieving an evaluation's purpose. Trevisan and Walser (2015) simplified Wholey's approach into a model of four iterative components:

1. focusing the EA,
2. developing an initial program theory,
3. gathering feedback on program theory, and
4. using the EA.

Each component features a checklist with questions reflecting the Program Evaluation Standards (Yarbrough et al., 2011).

The application of EAs has been intermittent. Although Wholey's work initiated a decade with a flurry of EAs in the 1970s–1980s, a revival in their use, particularly at the international level, did not occur until the late 1990s. Few publicly available examples of EAs exist. As Scriven (1991) warned about the confusion EAs may present in relation to other evaluative activities, reviews of the available EAs reflect concern about inconsistent implementation and use, revealing a lack of clarity about EA as a unique concept (Davies, 2013; Davies & Payne, 2015; Trevisan, 2007). Confusion with needs assessments, formative evaluations, and process evaluations were found, as was mission creep, with EAs extending into other evaluative functions based on commissioners' interests and budgets. Davies and Payne (2015) identified a need for clearly bounded expectations of the outputs of an EA and for linking the contents of a checklist for implementing EAs to relevant wider theory and evidence.

In developing our EA approach to CCA programming at the portfolio level, we found few concrete examples of previous EAs, with the exception of the Green Climate Fund Independent Evaluation Unit's 2019 Summary of the Evaluability of Green Climate Fund Proposals (Fiala et al., 2019). Given the challenges noted in the literature on maintaining clear objectives for EAs, and a need to have clear links between the dimensions of our checklist and relevant theory and evidence, we were purposeful in developing our EA framework, discussed below.

Evaluation of Climate Change Adaptation

Climate change poses dire consequences for the world, and given the relatively short timeframe to reverse trends, evaluating CCA interventions is essential to understand how best to adapt. The challenges in evaluating CCA interventions are well known and prove to be particularly complex. These include, for example, assessing attribution, creating baselines, and monitoring over longer time horizons (Bours et al., 2014; Fisher et al., 2015; Uitto et al., 2017).

CCA programming poses particular challenges for evaluation due to adaptation performance stretching far beyond the project life cycle. As a result, the impacts of such programs are difficult to measure because their outcomes may manifest much later. Climate change patterns and the prediction of weather patterns and extremes also pose a level of uncertainty for both programming and evaluation. The level of uncertainty increases when moving from global to regional climate models, to regional scenarios, and then to local impacts on human and natural systems (Wilby & Dessai, 2010). Given this unpredictability, collecting baseline data against which progress can be tracked is difficult. With climate change interventions spanning sectors and areas, another challenge is thinking at a systems level and across multiple stakeholder groups.

Finally, like most complex problems, CCA presents a challenge for causal inferences

between intervention and outcome. Given the cross-sector nature of CCA, with multiple influences from both the social and natural worlds, developing a coherent evidence base on which to make causal inferences is difficult (Bours et al., 2015). In spite of these challenges, MEL for CCA projects plays a central role for identifying how best to reduce vulnerability and build resilience to climate change (Bours et al., 2014). With a growing need for accountability and learning, having MEL systems in place that generate evidence that is fed back into adaptation practice is important. In this context, the role of evaluability is critical to ensuring that a project can be evaluated and has the foundations necessary for carrying out evaluations that will offer important lessons for the future.

Study Objectives

Given the need for clearer definition on EAs and the challenges CCA programming presents for evaluation, this chapter describes the Fund's evaluability framework and the process of developing and applying it in assessing the evaluability of the projects part of the Fund's portfolio. We reflect on areas of learning that have implications for both the evaluation and CCA fields.

Assessment Approach

Framework Development

The AF-TERG embarked on an EA in 2019 to examine all 100 projects approved by the Fund's board at the start of the assessment's inception in November 2019, making up the Fund's project portfolio. These projects were diverse, spanning a range of types of grantee organizations working within varying contexts, from small island states in the Caribbean and Pacific to locations in Central Asia, Africa, and South America. The projects also had a diverse set of implementing entities and stakeholders, including grassroots organizations, civil society organizations (CSOs),

government bodies, regional organizations, and multilateral stakeholders.

We adopted the OECD DAC (2002) definition of evaluability, "the extent to which an activity or a program can be evaluated in a reliable and credible fashion," and developed two objectives to guide the development and implementation of our EA:

1. Assess the extent to which the Adaptation Fund's projects have in place structures, processes, and resources capable of supporting credible and useful monitoring, evaluation and learning.
2. Based on the assessment's findings, provide advice on how to improve the evaluability of the Adaptation Fund's projects and portfolio.

The bounded outputs for the assessment included determining the extent of evaluability of the overall portfolio and identifying ways to improve evaluability. These findings were discussed internally to develop strategies to address both policy and operations, and to inform future evaluative work. Our EA design was guided by Scriven's (2007) discussion on checklists, an approach that implies a comprehensive approach to understanding the phenomenon under study. Practical considerations for implementation, including review of the MEL standards already applied by the Fund and standards not necessarily applied but identified as being of critical importance to credible MEL and/or evaluability, were based on the work of Davies (2013) and Davies and Payne (2015). The design was conceptualized in two phases, assessment and verification. The first phase, which we detail below, was concluded by July 2020, with the second phase of field verification planned for 2021.

Through a process of literature review, brainstorming, and multiple consultations, we constructed our assessment framework with seven categories, each associated to a key component of MEL: (a) project logic, (b) MEL plan and resources, (c) data and methods, (d) inclusion, (e) portfolio alignment, (f) long(er)-term evaluability, and (g) evaluability in practice. The implication for these multiple categories as a checklist is

that, taken together, they make up the totality of what the AF-TERG identified as evaluable CCA projects. Within each category, we established a series of assessment criteria. The categories and criteria were a combination of (a) MEL standards already applied by the Adaptation Fund, and (b) standards not necessarily applied by the Fund but identified by the AF-TERG assessment team as standards and approaches of critical importance to credible MEL and/or evaluability, particularly for CCA-focused projects. Table 1 presents the categories, their criteria, and brief discussion of their relevance.

Process for Implementation

As part of the first phase, we reviewed the original proposal documentation for all 100 board-approved projects.¹ Proposal documentation was the main documentation for analysis in the sense of (a) making a consistent assessment of the evaluability of the portfolio at the project onset and of the structures already available and considered for evaluation in the project design, and (b) informing the MEL and proposal review processes in terms of evaluability in the Fund. We also analyzed available project inception reports, project performance reports, midterm evaluation reports, and terminal evaluations to develop an initial understanding of the evaluability in practice during project implementation.

Consistent with the intentions that the objectives of the assessment were clearly laid out and that the team would carry out a second phase of field verification for evaluability in practice, the first phase was exclusively desk based. The nature of the phased approach restricted the depth of analysis: Where gaps or uncertainties occurred within individual projects, we did not seek clarification through means such as follow-up interviews with implementing entities or project teams. However, limiting the work exclusively to a desk review allowed for broad coverage and

enabled analysis of the entire portfolio, while also shedding light on the Fund's existing review process. The second phase was intended to fill any information gaps identified during the first phase.

Based on a detailed set of guidance, reviewers assessed the evaluability of all 100 projects against each of the assessment criteria, providing narrative justification for assessments and allocating ratings where relevant (where logical, we applied rating scales to individual criteria). Table 2 provides an example of a criterion's rating system. All project-level assessments were recorded within a spreadsheet-based assessment tool that, in turn, supported portfolio-level analyses and in the longer term will serve as a transparent, accessible database of all the EA's underlying data.

In addition to the seven assessment categories and their underlying criteria, the tool also recorded descriptive detail for every project, such as project status, budget, country, and sector(s), for example. We subsequently used this detail to support cross-portfolio analyses, enabling the assessment to identify patterns or trends in evaluability according to criteria such as a project's age, the kind of implementing entity that leads the project, the context within which the project works, and so forth.

Analysis

The results of the portfolio evaluability assessment illustrate the common tensions already encountered in CCA monitoring and evaluation (M&E). More precisely, they concur with the three challenges identified by Fisher et al. (2015): assessing attribution, creating baselines, and monitoring climate change activities over longer time horizons. The assessment of various criteria allows us to pinpoint specific elements in the project design and implementation that contribute to these tensions. Finally, the results show the importance of undertaking evaluability assessments to help build stronger M&E in the field of CCA.

¹ The projects were at different points in the project cycle: approved but implementation had not started, under implementation, or completed at the time of assessment.

Table 1 The Adaptation Fund's evaluability assessment framework

Category	Project logic			
Criteria	Quality of project logic	Quality of evidence base	Clarity of project additionality	
Project logic is integral to project evaluability because it describes the basis and justification for an intervention, the starting conditions and assumptions, the expected results, and the anticipated means or pathways through which the project will deliver those results. CCA-focused projects should also be based on a logic model that reflects—and identifies linkages between—both the human and natural systems being affected.				
Category	MEL plan and resources			
Criteria	Quality of M&E plan	Quality of MEL resources	Quality of approach to learning	
For a project to be evaluable, there must be a feasible approach to identifying the project’s likely contribution to outcomes, including a clear direction as to what aspects of the project will be measured and how, a description of who will be involved (roles and responsibilities), the resources that will be available to deliver the proposed approach, and intended use for adaptive management.				
Category	Data and methods			
Criteria	Quality of results statements	Quality of indicators or other measures	Quality of baselines	Quality of data
Project evaluability is often dependent on the relevance and quality of data generated during project implementation, with the gathered data typically (but not necessarily exclusively) defined by a project’s indicators. Consequently, an assessment of project evaluability needs to consider the quality of project indicators, the methods through which progress against those indicators is assessed, and the quality of baseline against which project performance and results will be assessed.				
Category	Inclusion			
Criteria	Quality of project logic	Quality of evidence base	Clarity of project additionality	
	Quality of data disaggregation: gender		Quality of data disaggregation: other	
The assessment considered the extent to which MEL activities involved and represented all of a project’s interest groups, whether individuals, communities, sub-national institutions, or country governments. This includes assessment of the participation and representation of women, youth, and socially excluded and vulnerable groups. Recognition is growing that an intervention’s approach to MEL—and the data that MEL generates—can only be credible if it is based on the consent and participation of the people and institutions that the intervention aims to support. This is particularly the case for CCA, where interventions are frequently focused on building individual, community, and institutional resilience.				
Category	Portfolio alignment			
Criteria	Depth of alignment		Quality of monitoring and reporting against portfolio results	
Data and learning generated through each project’s MEL activity should also contribute to and strengthen the evaluability of the Adaptation Fund’s overall portfolio. Project-level contributions to portfolio evaluability will be most readily achieved where clear alignment exists between a project’s expected results and the Adaptation Fund’s Strategic Results Framework.				

(continued)

Table 1 (continued)

Category	Long(er)-term evaluability	
Criteria	Potential for postcompletion evaluation	
The kind of results and changes targeted by CCA interventions are mostly long(er)-term in nature, only identifiable and measurable well after a project has been implemented. But for an evaluation to be undertaken 5 years after a project’s funding period, for example, resources must be available and—more important—sufficient foundations must be in place for an evaluation to be even plausible. The prospects for long(er)-term evaluability could be improved by, for example, a project having indicators that can be accurately measured over longer time horizons, a project logic model that extends beyond implementation, and the likely availability of institutions and individuals that participated in the original intervention.		
Category	Evaluability in practice	
Criteria	Changes to MEL approach documented and justified	Quality of evaluability in practice
Documentation such as midterm and terminal evaluations can reveal how well a project’s MEL strategy has performed in practice and, by extension, can provide insight into the actual strengths and challenges around project evaluability. Consequently, analyzing this documentation can improve understanding of the practical limitations (and opportunities) that projects face when it comes to ensuring interventions are evaluable. Comparing these documents to a project’s original MEL design can also build understanding as to whether and how MEL strategies typically change. Once a project is under implementation, adjustments to MEL strategy are often required due to situations such as changes in plans, unanticipated monitoring challenges, or unforeseen resource constraints. Under such circumstances, project evaluability can be maintained and improved if changes to MEL strategy are well documented within progress reports.		

Table 2 Example of a criterion’s rating system

Criterion	Quality of project logic					
Guiding question	Does the project logic provide a sufficient basis against which performance and results can be evaluated?					
Scale	Very good quality	Good quality	Fair quality	Weak quality	Very weak quality	Logic not documented
Guidance	<p>A project’s logic is considered very good quality if it includes all of the following:</p> <ul style="list-style-type: none">Explicit logic model, whether a theory of change, logical framework, or other explicit modelClear description of the project’s expected resultsClear description of the pathways and processes through which the project will deliver expected resultsClear description of the assumptions underpinning the project’s logicClear description of the external influences that may affect project deliveryIntervention logic expresses a clear contribution to CCAProject logic reflects both human and natural systemsEnvironmental and social risks are reflected in project logic <p>The scale should then be applied according to the degree of alignment with the above elements and the clarity of discussion around project logic, with—at the lower end of the scale—very weak quality meeting only one of the above elements and/or logic is unclear.</p>					

Logic and Additionality of Adaptation Projects

Relevance to Evaluability

To be evaluated effectively, a project must have clearly articulated logic. Clarity on additionality also strengthens evaluability by improving the potential for identifying (isolating) a project’s influence and results. Within the project logic

evaluability category, the assessment looked at the criterion clarity of project additionality (see Table 3).

Adaptation-Specific Evaluability Considerations

CCA projects invariably work within complex natural systems and target results that focus on human interactions within those systems. Those

Table 3 Guidance for assessing clarity of project additionality

Category	Project logic	
Criterion	Clarity of project additionality	Guiding question for assessors
		To what extent is it clear that the project’s work is additional to business-as-usual and/or other initiatives that have been or are being delivered?

complex systems are influenced by multiple factors, many (if not most) of which will be completely outside the control of any given intervention. Similarly, to *attribute* higher-level results to a single project is often unrealistic or erroneous. The additionality of a CCA project is more often described in terms of the *contribution* to higher level results and the project’s impact pathways toward those results. From an evaluability perspective, this necessitates a clear description not just of a project’s additional contribution, but also the numerous other factors that will influence higher level results, and a project’s interactions with those other factors.

Findings

The assessment found that clarity of project additionality was greatly supported by a project proposal template that required Fund applicants to provide detail on related initiatives and describe whether and how their proposed work duplicated those other initiatives. This requirement was primarily motivated by the Adaptation Fund’s desire to ensure that their grants did not duplicate other funding sources. However, this simple, standard request also helped evaluability: It obliged applicants to think through and articulate how their project related to broader work on adaptation and, in doing so, helped to ensure that project proposals invariably contained clear descriptions of how interventions were (or were not) additional to external initiatives.

However, from an evaluability perspective, this created a common tension. Projects understandably and rightly sought to gain efficiencies and synergies with related interventions, often to the extent that the clarity of additionality was

reduced. Where the assessment identified a lack of clarity around an intervention’s additionality, this was often because of programmatic strengths and efforts to contribute to a broader agenda. But when resources and efforts were combined across projects, perceiving how any eventual results could be directly attributable to a specific intervention or individual funding source became more difficult; at best, only a project’s contribution to results would be evaluable.

As noted above, this scenario of multiple actors and influences working within complex systems is common (even standard) for CCA projects. The evaluability of such projects will benefit from clear, honest descriptions of how a project is distinct and—just as important—how a project intersects or even duplicates other work. Given that interdependencies between CCA projects are common, evaluability may be reliant on description and analysis of not just the main project’s own results chains, but also the results chains of related initiatives. Furthermore, projects—and the funders that support those projects—may need to accept that the only measurement ever possible for a project may be its contribution to high-level results.

Evidence Base and Baselines: Natural vs. Human Systems

Relevance to Evaluability

When it comes to a project’s evidence base, evaluability can be strengthened where there are clear lines and logical linkages between prior experiences and a proposed intervention. The evidence base and learning from previous work can also help to define how performance and results should be measured, including the design and setting of baselines. In turn, baselines are critical for evaluability: A project’s progress can only be monitored, evaluated, and fully understood if comparisons can be made between a project’s current position and a clearly described starting point that a baseline ideally establishes. The assessment looked at quality of evidence base and quality of baseline across two evaluability categories, project logic and data and methods (see Table 4).

Table 4 Guidance for assessing quality of evidence base and baseline

Category	Project logic	
Criterion	Quality of evidence base	Guiding question for assessors: Does the project logic reference and take into account data, evidence and learning from prior research, initiatives, and institutional experience?
Category	Data and methods	
Criterion	Quality of baseline	Guiding question for assessors: Does a baseline exist for all of the proposed project indicators/measures, or are clear plans in place for developing a baseline?

Adaptation-Specific Evaluability Considerations

Of particular relevance to CCA-focused work and to Adaptation Fund projects specifically, the assessment looked at how project evidence base and baselines took into account both natural and human (including institutional) systems. The conceptualization of—and separate emphasis on—natural and human systems is central to the Fund’s definition (AF, 2017, p. 3) of an adaptation project:

A concrete adaptation project/programme is defined as a set of activities aimed at addressing the adverse impacts of and risks posed by climate change. The activities shall aim at producing visible and tangible results on the ground by reducing vulnerability and increasing the adaptive capacity of human and natural systems to respond to the impacts of climate change, including climate variability.

A CCA project’s evaluability will be partly determined by how well a project’s evidence base and baseline describe the starting position of the targeted natural systems (e.g., forestry coverage, biodiversity, soil characteristics), the starting position of the targeted human systems (e.g., agricultural practices, economic incentives, government policy, institutions), and the current ways in which the two systems interact (e.g., agricultural practice reducing biodiversity, economic incentives accelerating deforestation).

Findings

With regard to additionality, the Adaptation Fund’s proposal templates ensured that projects, in the main, presented clear, well-referenced descriptions of the preintervention evidence base. This requirement was principally used by Fund applicants to justify the case for a project, but it also served to strengthen evaluability: The evidence base and site-specific context inherently provided a basis against which project progress and results could be evaluated.

However, the assessment also found that the depth and quality of the preintervention evidence base varied according to whether evidence related to natural systems or to human systems. The use of detailed climatic and environmental baseline data (i.e., natural systems) was especially strong within most proposals. Conversely, proposals devoted less attention to describing and evidencing the nonenvironmental context within which an intervention was to operate, including the human and institutional aspects of a project. Beginning at the design/proposal stage, evaluability tended to be stronger for work relating to natural systems, less so for work relating to human systems (and, by extension, the linkages and interactions between natural and human systems).

While considering how the preintervention evidence base and baselines reflected human and natural systems, the evaluability assessment also reviewed the extent to which—once projects were under implementation—monitoring approaches (including results frameworks) measured change across both systems, and the interdependencies between them. Although the majority of projects did support some degree of measurement of both human and natural systems, direct monitoring approaches were very heavily geared toward only measuring change within human systems (e.g., agricultural infrastructure, institutional and individual capacities, legislation). Indeed, many projects that measured some aspect of natural systems did so through only one relatively high-level indicator (e.g., area of land restored). Moreover, only a handful of projects had results frameworks in place that would be capable of measuring both systems and the interdependencies between them. The comparatively strong baseline understanding of natural systems

was often not being expanded (or even followed up) during implementation. Perhaps this is because accessing historical natural data (i.e., climate, environment, biodiversity) is comparatively easy, so a strong evidence base can be developed. Conversely, accessing historical human data could be more difficult (and such data may not be available), thus developing a detailed pre-implementation evidence base could be quite resource intensive.

Aside from the differing treatments of human and natural systems, the findings also illustrate that ensuring project evaluability is an ongoing process. Evaluability can’t be achieved through project design alone: Even where evaluability appears strong at the pre-implementation stage, that level of evaluability has to be maintained throughout project delivery by, for example, designing and applying processes capable of gathering the breadth and quality of data necessary to measure results across both human and natural systems.

Resources Allocated to MEL: Direct vs. Indirect

Relevance to Evaluability

Project evaluability is partly dependent on sufficient institutional and/or financial resources being allocated toward MEL activity (with the definition of what comprises “sufficient” resources dependent on the nature of each project and its MEL strategy; see Table 5). To support the assessment of resource adequacy, we identified the level of financial resources allocated toward MEL for every project, recording two figures:

- Direct resources: Money allocated explicitly toward MEL activities.
- Indirect resources: Money allocated towards activities that, although not itemized or categorized as MEL, are likely to be of direct, substantive benefit to MEL.

Adaptation-Specific Evaluability Considerations

Many CCA interventions include—and are sometimes focused exclusively on—activity that

Table 5 Guidance for assessing quality of MEL resources

Category	MEL plan and resources	
Criterion	Quality of MEL resources	Guiding question for assessors:
		Are financial and institutional resources for MEL explicitly defined, and are these sufficient to support delivery of the MEL strategy?

can be considered indirect MEL. Examples of such activities are development of new climate monitoring approaches, consolidation of historical environmental data, research activities, and capacity development on the monitoring and interpretation of data. Project evaluability can be strengthened where these indirect activities are formally linked with the intervention’s MEL strategy, in turn ensuring that project MEL (and evaluability) benefits from the widest possible range of data and resources.

Findings

During assessment of this criterion, our most notable finding was that the proportion of indirect financial resources associated with MEL-relevant activity was far larger than the proportion of resources directly allocated to MEL. Frequently, indirect MEL activity was a core component of projects, such as in projects focused almost exclusively on developing sub-national climatic monitoring infrastructure. Many projects also were strengthening institutions, capacities, and infrastructure required for longer term (postproject) monitoring and measurement of CCA-relevant indicators and results.

However, such obviously MEL-relevant activities were not recognized or presented as such by projects. This disconnect could have been due to project partners not conceptualizing the activity as MEL, or the common perception of MEL as a purely accountability-focused exercise, rather than as a tool that can also support learning and knowledge generation. Regardless of the reasons for the disconnect, where projects did not make the link between their direct MEL work and those indirect activities with clear relevance and value for MEL, the data and learning generated through

indirect activities may have provided only a limited contribution or none at all to the project’s formal MEL. This could have reduced the potential and strength of a project’s evaluability.

Another notable finding was that the data collected through indirect MEL activities tended to be more focused on natural than on human systems. While some projects were actually gathering extensive data and/or developing infrastructure for the measurement of natural systems, this data was not being used to bridge the above-noted gap whereby formal, direct project MEL placed more (and sometimes exclusive) emphasis on the monitoring of human systems. Thinking beyond individual projects, we also see a risk that data could be overlooked by future research and meta-evaluations if not identified as being MEL relevant. In turn, this could reduce the potential contribution and value of data to longer-term learning and knowledge generation.

From an evaluability perspective, these findings highlight the importance of looking beyond a project’s self-identified direct MEL activity, particularly where those projects are working to develop monitoring capacities and infrastructure. Moreover, the findings also demonstrate how the actual process of evaluability assessment can help to identify and uncover opportunities for strengthening not just evaluability, but a project’s MEL more broadly.

Potential for Postcompletion Evaluation

Relevance to Evaluability

Although still a relatively underdeveloped approach, postcompletion evaluations are increasingly deployed as a tool to measure the longer term impact of interventions. However, for an evaluation to be undertaken several years after a project’s closure, adequate resources must be available and, more important, sufficient foundations must be in place for an evaluation to be even plausible (see Table 6). The prospects for long(er)-term evaluability could be improved by, for example, a project having indicators that can be accurately measured over longer time horizons, a project logic model that extends beyond

Table 6 Guidance for assessing potential for postcompletion MEL

Category	Long(er)-term evaluability	
Criterion	Potential for postcompletion MEL	Guiding question for assessors:
		Are the project’s logic, monitoring approach, and postcompletion institutional arrangements sufficiently long(er)-term in nature to support postcompletion MEL?

implementation, and the likely longer term availability of institutions and individuals that participated in the original intervention.

Adaptation-Specific Evaluability Considerations

The kind of results and changes targeted by CCA interventions are mostly longer term in nature, only identifiable and measurable well after a project has been implemented. Consequently, the justification for postcompletion evaluation is particularly strong in the adaptation arena. Therefore, considering the extent to which foundations are in place for longer-term evaluability—and postcompletion evaluation—is important for any assessment of an adaptation project’s evaluability.

Findings

The assessment first sought to identify whether projects had formal plans in place for postcompletion evaluation. Only 3% of projects confirmed formal plans for postcompletion evaluations, with even the concept of postcompletion evaluation rarely being mentioned across the portfolio. Plans for postcompletion monitoring were far more prevalent, with several projects including components whose entire purpose was to establish organizational structures, systems, and capacities for longer-term (indefinite) monitoring of factors such as local meteorological data, water levels, or land use. However, the funding and institutional arrangements for these longer term systems were infrequently specified. All of these longer term monitoring systems focused on only one aspect of

a project; we identified no examples of projects that planned broader longer term monitoring through approaches such as continued, wide-ranging monitoring against their original results framework.

The limited examples of longer term MEL are unsurprising; even within the MEL sector the concept of longer term MEL is still relatively new. Consequently, the assessment also considered the extent to which projects had in place certain foundations that at least strengthened the potential for postcompletion MEL, such as logic models that established pathways and results beyond the project's lifetime, project indicators that could plausibly be measured 5 years after project completion, indicators based on preexisting national data sets that are likely to be maintained over the longer term, and clear descriptions of postcompletion institutional ownership of the project's outcomes. Some instances of projects had promising building blocks in place, with several projects benefiting from clear exit strategies and descriptions of postcompletion ownership, whether institutional, community, or individual. A handful of projects also aligned their results frameworks with preexisting national results frameworks and data sets, with the explicit rationale being to support longer term monitoring of project results. However, these were exceptions; the broader portfolio was characterized by generally weak potential for postcompletion MEL.

Again, these findings are not surprising given the universally low application of longer term MEL and postcompletion evaluation. However, the process of evaluability assessment can help pinpoint gaps and opportunities to at least strengthen the foundations for any potential postcompletion evaluation.

Reflections on the EA Tool Development and Implementation

We undertook a reflective and consultative process to develop a tool specific to CCA programming and its particular challenges, aiming to fulfill our objective of an impartial and accurate assessment of the Adaptation Fund portfolio's evaluability. The tool was informed by a set of

principles taking into account the Fund's current and historical approaches to evaluation and an understanding of our evaluand, CCA programs. We adopted the OECD definition of evaluability and set about a process akin to both Wholey's (1979) eight steps and Trevisan and Walser's (2015) four-stage approach: defining the evaluand, literature review, framework development, consultation and piloting, implementation, and presentation of results.

The process of our EA tool development was consultative, involving peer review by an evaluation consultant and by AF-TERG's advisory board of evaluation and climate change programming experts. We developed a 5-point scale with criteria of equal weighting. We piloted the tool, reflected on its results, further revised, and piloted again. Two colleagues undertook the review, aligning their interpretation of criteria and scoring. Although we did not determine an inter-rater reliability score, we systematically compared assessments to calibrate judgment.

The results of the EA identified challenges that are well documented about CCA programming (Fisher et al., 2015), as discussed above, and resonated with Adaptation Fund stakeholders, serving to promote discussion and decision making about the Fund's evaluation policy and processes for funding and partnership. A consideration for this first phase of the EA was the sole reliance on written documentation and the lack of opportunity to engage on site with partner agencies. In the second phase of our two-phased approach of assessment and verification, we will continue to collect data on site with partners and reflect on the tool's merit and utility as part of our effort to engage in formative meta-evaluation. With this positive EA experience and as we find additional use for the tool going forward, we may also meta-evaluate following field verification, using relevant criteria found in the Program Evaluation Standards (Yarbrough et al., 2011).

Conclusions

The framework that was developed and applied for the Adaptation Fund EA proved to be a useful tool for understanding evaluability strengths and

gaps both within individual projects and across the Fund's whole portfolio. Moreover, the framework supported the identification of CCA-specific evaluability opportunities and challenges. Results of the assessment supported the three challenges related to climate change activities identified by Fisher et al. (2015): assessing attribution, creating baselines, and monitoring over longer time horizons.

The assessment confirmed the difficulty in identifying the additionality of CCA programs because of the complexity of such interventions. Clarifying the project logic through an EA helps conceptualize the project results in terms of contribution or attribution in a longer term perspective.

The study also highlighted the difficulty in setting baselines to measure the results of CCA in a comprehensive way, which would encompass impacts on both natural and human systems. It showed that, beyond the usual suspect of unavailability of nature-focused and human-focused data, the failure to integrate this data from both systems in the MEL system was the source of gaps in knowledge during project implementation.

Finally, the study highlighted the need to plan for and allocate resources for longer term M&E to be able to measure the delayed impacts of CCA interventions. The EA process can help pinpoint gaps and opportunities to at least strengthen the foundations for any potential postcompletion evaluation. The EA was incredibly useful in assessing the extent of the structures, processes, and resources in place capable of supporting credible M&E. Although we conducted this assessment retroactively, prospective use of the tool could inform and ensure that projects have the evaluability structures necessary at project design. The study of evaluability in practice, which verifies whether the evaluability is maintained throughout project delivery, could also be complemented with field verification to inform gaps in the processes related to project implementation.

The framework also offered a structured process for systematically thinking through what constitutes sound MEL more broadly. Indeed, a key conclusion of the assessment was that the

strength and quality of a project's overall approach to MEL is a major determinant of a project's evaluability: If a project gets its MEL "right," the project is highly likely to also have strong evaluability. This suggests that, where improved evaluability is sought, focusing efforts specifically on strengthening evaluability may not be efficient or even necessary; instead, improvements to overall MEL strategy and processes should inherently deliver improvements to the quality of evaluability.

References

- Adaptation Fund. (2011a). *Guidelines for project/programme final evaluations*. Author. <https://www.adaptation-fund.org/document/guidelines-for-projectprogramme-final-evaluations/>
- Adaptation Fund. (2011b). *Results framework and baseline guidance – Project-level*. Author. <https://www.adaptation-fund.org/wp-content/uploads/2015/01/Results%20Framework%20and%20Baseline%20Guidance%20final%20compressed.pdf>
- Adaptation Fund. (2012). *Evaluation framework*. Author. <https://www.adaptation-fund.org/document/evaluation-framework-4/>
- Adaptation Fund. (2017). *Operational policies and guidelines for parties to access resources from the adaptation fund*. Author. <https://www.adaptation-fund.org/document/operational-policies-guidelines-parties-access-resources-adaptation-fund/>
- Bours, D., McGinn, C., & Pringle, P. (2014). *Monitoring & evaluation for climate change adaptation and resilience: A synthesis of tools, frameworks and approaches* (2nd ed.). SEA Change Community of Practice and UKCIP. <https://ukcip.ouce.ox.ac.uk/wp-content/PDFs/SEA-Change-UKCIP-MandE-review-2nd-edition.pdf>
- Bours, D., McGinn, C., & Pringle, P. (2015). *Monitoring and evaluation of climate change adaptation: A review of the landscape*. Wiley.
- Davies, R. (2013). *Planning evaluability assessments: A synthesis of the literature with recommendations* (Working Paper 40). Department for International Development London, United Kingdom. <https://www.gov.uk/dfid-research-outputs/planning-evaluability-assessments-a-synthesis-of-the-literature-with-recommendations-dfid-working-paper-40>
- Davies, R., & Payne, L. (2015). Evaluability assessments: Reflections on a review of the literature. *Evaluation*, 21(2), 216–231. <https://doi.org/10.1177/1356389015577465>
- Fiala, N., Puri, J., & Mwandiri, P. (2019). *Becoming bigger, better, smarter: A summary of the evaluability of Green Climate Fund proposals*. Green Climate Fund

- (GCF) Independent Evaluation Unit (IEU). https://ieu.greenclimate.fund/documents/977793/985626/Working_Paper__Becoming_bigger__better__smarter_-_A_summary_of_the_evaluability_of_GCF_proposals.pdf
- Fisher, S., Dinshaw, A., McGray, H., Rai, N., & Schaar, J. (2015). Evaluating climate change adaptation: Learning from methods in international development. In D. Bours, C. McGinn, & P. Pringle (Eds.), *Monitoring and evaluation of climate change adaptation: A review of the landscape* (pp. 13–35). Wiley.
- Organisation for Economic Co-operation and Development, Development Assistance Committee. (2002). *Glossary of key terms in evaluation and results based management*. Author. <https://www.oecd.org/dac/2754804.pdf>
- Rutman, L. (1980). *Planning useful evaluations: Evaluability assessment*. Sage.
- Scriven, M. I. (1991). *Evaluation thesaurus* (4th ed.). Sage.
- Scriven, M. I. (2007). *The logic and methodology of check-lists* [unpublished manuscript]. Western Michigan University.
- Smith, M. (1989). *Evaluability assessment: A practical approach*. Kluwer Academic.
- Stufflebeam, D., & Coryn, C. (2014). *Evaluation theory, models, and applications* (2nd ed.). Jossey-Bass.
- Trevisan, M. S. (2007). Evaluability assessment from 1986 to 2006. *American Journal of Evaluation*, 28(3), 290–303. <https://doi.org/10.1177/1098214007304589>
- Trevisan, M. S., & Walser, T. M. (2015). *Evaluability assessment: Improving evaluation quality and use*. SAGE.
- Wholey, J. S. (1979). *Evaluation: Promise and performance*. Urban Institute.
- Wholey, J., Hatry, H. P., & Newcomer, K. E. (2010). *Handbook of practical program evaluation* (3rd ed.). Jossey-Bass.
- Wilby, R. L., & Dessai, S. (2010). Robust adaptation to climate change. *Weather*, 65(7), 180–185. <https://doi.org/10.1002/wea.543>
- Uitto, J. I., van den Berg, R. D., & Puri, J. (2017). *Evaluating climate change action for sustainable development*. Springer Open.
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Sage.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.