



ADAPTATION FUND

AFB/EFC.34/Inf.6
2 October 2024

Adaptation Fund Board
Ethics and Finance Committee
Thirty-fourth Meeting
Bonn, Germany, 8 - 9 October 2024

Agenda Item: 5 (d)

Second Synthesis of Adaptation Fund Final Evaluations

FINAL REPORT

Synthesis of Adaptation Fund Final Evaluations (July 2020 to August 2023)

22 March 2024

Table of Contents

List of figures.....	iv
List of tables.....	iv
Acronyms.....	v
Executive summary.....	1
Quality of the final evaluation reports.....	2
Quality of the projects/programmes.....	3
Programmatic challenges and issues.....	3
Lessons from the portfolio.....	4
Recommendations.....	4
Programmatic recommendations for the AF Secretariat.....	4
1. Introduction.....	1
2. Background to the evaluation synthesis.....	1
3. Purpose, objectives, and scope of the FE synthesis.....	2
Purpose.....	2
Objectives.....	2
Scope.....	2
4. Key synthesis questions.....	3
5. How to read this report.....	3
6. Methodology.....	3
Limitations.....	7
7. Portfolio overview.....	7
8. Quality of final evaluation reports.....	12
Dimension 1: Structure and clarity of the report.....	13
Dimension 2: Context, purpose, scope, and objective.....	14
Dimension 3: Evaluation framework and methodology.....	15
Dimension 4: Evaluation methods and data.....	17
Dimension 5: Analysis and findings.....	18
Dimension 6: Conclusions, lessons, recommendations.....	19
Dimension 7: Inclusion, ethics, independence.....	20
Dimension 8: Management and governance.....	21
Dimension 9: Utility.....	22
Section conclusion.....	23
9. Quality of projects/programmes.....	24

Effectiveness	25
Likelihood of impact.....	25
Estimated likelihood of sustainability.....	26
Section conclusion	26
10. Lessons from the synthesis	27
Programmatic challenges and issues	27
Lessons from the portfolio.....	30
11. Recommendations	31
Programmatic recommendations for the AF Secretariat	31
Annexes.....	33
Annex 1 Terms of reference.....	34
1. Background	34
3. Purpose and objectives of the second final evaluation synthesis	36
4. Methodology.....	36
5. Organization and schedule	38
Annex 2 Final evaluation quality review framework	40
Annex 3 Intended outcomes of projects/programmes covered by this synthesis	48
Annex 4 Confidential annex on the details of the evaluations and IEs....	Error! Bookmark not defined.

List of figures

Figure 1 Four-step methodology of the synthesis.....	4
Figure 2 Implementing entities	8
Figure 3 Aggregate AF Grant to projects/programmes evaluated by FEs within the scope of this synthesis, by region.....	9
Figure 4 Sector classification of the projects/programmes in the synthesis portfolio	10
Figure 5 Implementation period of the projects/programmes.....	10
Figure 6 Planned versus actual closing date of projects/programmes	11
Figure 7 Year of completion	12
Figure 8 Evaluation completion year and Figure 9 Evaluations undertaken during the pandemic.....	13
Figure 10 Ratings for Dimension 1 Structure and Clarity	14
Figure 11 Ratings for Dimension 2 Context, purpose, scope, and objective	15
Figure 12 Ratings for Dimension 3 Evaluation framework and methodology	16
Figure 13 Methodology adopted	16
Figure 14 Ratings for Dimension 4 Evaluation methods and data	17
Figure 15 Methods used in the evaluations.....	17
Figure 16 Ratings for Dimension 5 Analysis and findings.....	18
Figure 17 Ratings for Dimension 6 Conclusions, lessons, recommendations.....	20
Figure 18 Ratings for Dimension 7 Inclusion, ethics, independence	21
Figure 19 Ratings for Dimension 8 Management and governance.....	22
Figure 20 Ratings for Dimension 9 Utility	23
Figure 21 Overall quality rating of the final evaluation reports.....	24
Figure 22 Project/programme effectiveness ratings from the two syntheses.....	25
Figure 23 Project/programme impact ratings from the two syntheses.....	26
Figure 24 Estimated project/programme likelihood sustainability ratings from the two syntheses	26
Figure 25 Overall project/programme ratings from the two syntheses	27
Figure 26 Relevance ratings from the first and second synthesis	28
Figure 27 M&E ratings from the first and second synthesis	29
Figure 28 Efficiency ratings from the first and second synthesis	29
Figure 29 Reasons for delays in project/programme implementation	30
Figure 30 Three-step approach to the development of the draft methodology	40

List of tables

Table 1 Mapping of report sections against the synthesis objectives.....	3
Table 2 Projects/programmes included in the synthesis	8
Table 3 AF Grant received by the projects/programmes evaluated by FEs within the scope of this synthesis, by country.....	9
Table 4 Project/programme start-up delays	10
Table 5 Magnitude of extension in the implementation period of the projects/programmes.....	11
Table 6 Rating system for each quality criterion	12
Table 7 Overall rating guide.....	23

Acronyms

AF	Adaptation Fund
AF-TERG	Adaptation Fund Technical Evaluation Reference Group
CCA	Climate change adaptation
DA	Designated Authorities
FAO	Food and Agriculture Organisation
GCF	Green Climate Fund
GEF	Global Environment Facility
IE	Implementing Entity
M&E	Monitoring and Evaluation
MIE	Multilateral Implementing Entity
MoE Rwanda	Ministry of Environment Rwanda
NIE	National Implementing Entity
OECD-DAC	Organisation for Economic Co-operation and Development's Development Assistance Committee
SANBI	South African National Biodiversity Institute
ToR	Terms of Reference
UK FCDO	United Kingdom Foreign, Commonwealth and Development Office
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme
UNFPA	United Nations Population Fund
WFP	World Food Programme

Executive summary

The Technical Evaluation Reference Group of the Adaptation Fund (AF-TERG) commissioned a short-term consultant to undertake a synthesis of the final evaluations submitted to the Adaptation Fund (AF) between July 2020 and August 2023. This report, the final deliverable of the consultancy, is the revised and final version of the draft evaluation synthesis report. It incorporates the feedback that AF-TERG provided between December 2023 and January 2024.

The AF Evaluation Policy and relevant guidance documents served as an anchor for this synthesis. However, the final evaluations covered in this synthesis were not yet subject to the application of the AF Evaluation Policy but were guided by the preceding guidelines, for example, the AF's Guidelines for Project/Programmes Final Evaluations and the AF's Evaluation Framework. As such, the quality review that this consultancy undertook were not intended to present the level of compliance of Implementing Entities.

This second synthesis builds on the AF-TERG's first synthesis. It aims to strengthen the AF-TERG's efforts to provide synthesis products that will strengthen learning within the AF and beyond. It will also inform the formulation and implementation of the AF-TERG's new multi-year work programme, further the work of the AF and its Implementing Entities (IE) and, above all, provide an important frame of reference for the AF Evaluation Policy.

The synthesis covers **12 final evaluations** implemented by a combination of Multilateral Implementing Entities (9 projects/programmes) and National Implementing Entities (3 projects/programmes) in Asia, Latin America and Caribbean, and Sub-Saharan Africa. It serves **learning, accountability**, and oversight purposes. The primary audience of the synthesis are the AF Board Secretariat and the AF Board. The secondary audience include the IEs, other climate funds and donors operating in the adaptation space, and other organizations and evaluators aiming to undertake similar syntheses.

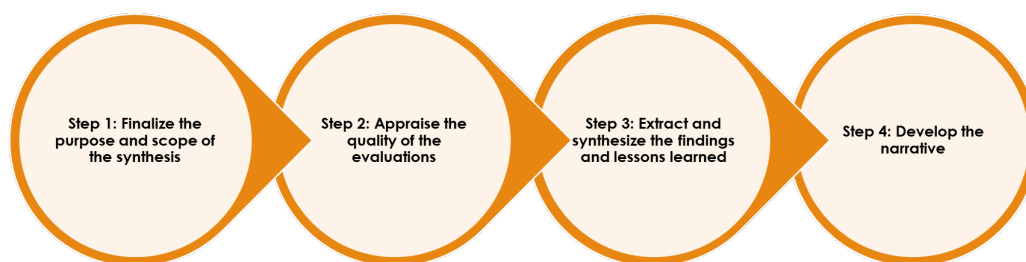
The synthesis has four main objectives:

1. To report to the Board on the quality of final evaluations and on the overall performance and effectiveness of completed projects and programmes, including in relation to improved adaptive capacity and resilience.
2. To review the quality of final evaluations reports based on compliance with the criteria as laid out in the Evaluation Policy and Guidance of Final Evaluations.
3. To assess the quality of the lessons in the final evaluations' recurring findings, lessons learned and evidence on specific thematic areas of interest such as gender/equity, improved adaptive capacity and resilience, or sustainability.
4. Serve an accountability purpose through exploring the extent of management response to recommendations provided in the Final Evaluations and how the management response is fed into future programming.

To respond to these objectives, the synthesis adopted a methodology that drew on the **combined strengths of realist synthesis and lessons learned synthesis**. This methodology was applied to the synthesis in four steps. The synthesis formulated and applied a **quality review framework** to assess the quality of the final evaluation reports. With nine quality dimensions and a 6-point Likert scale, overall scores were calculated using weights. The methodology was iterated twice following a pilot application to four evaluation reports. After the pilot, the remaining evaluation reports were assessed, and the findings analysed using descriptive and summative statistics.

To assess the quality of the projects/programmes, the synthesis considered the ratings for each evaluation criterion in the individual reports. They were then compared and visualised using Excel. Challenges and lessons learned were organized along thematic lines and included in the analysis where they were cited by

half of the reports (n=6) or more. Rapid thematic coding was used to generate these common challenges and lessons. Unsubstantiated lessons were excluded.



The **main limitation** of the synthesis is that it solely relied on secondary information in the 12 evaluation reports. There was no opportunity for verification and triangulation as well as further consultations to enrich the findings and recommendations.

Quality of the final evaluation reports

On balance, the quality of the portfolio of final evaluations included in this synthesis is satisfactory. The reports are largely compliant with the new AF Evaluation Policy and related guidance notes. Nine out of 12 reports were Moderately Satisfactory to Highly Satisfactory and only three were Unsatisfactory. Two of the three reports were commissioned by an NIE. Among the most common shortcomings from the perspective of the AF Evaluation Policy include incomplete evaluation matrix, unsubstantiated findings, lessons that were phrased as recommendations, limited lessons and recommendations beyond project/programme boundaries, unelaborated ethical standards, and management response that was not included in the report.

Dimension	Summary of findings
Dimension 1: Structure and Clarity of the Report	11 evaluation reports have moderate to no shortcomings hence scoring between Moderately Satisfactory to Highly Satisfactory. Evaluators either followed the outline included in the ToR or adopted the typical structure of an evaluation report. Reports are generally accessible to the readers and the narratives are easy to follow.
Dimension 2: Context, Purpose, Scope, and Objective	10 evaluation reports scored Highly Satisfactory. The reports normally situate the problem the project/programme is trying to address within context that has been described comprehensively. The context, purpose, scope, and objective are often extracted from the project documents in the case of context, and in the evaluation ToR in the case of purpose, scope, and objective.
Dimension 3: Evaluation Framework and Methodology	11 evaluation reports clearly specified and described the evaluation methodology and framework. They included either an evaluation matrix in the annex or a reconstructed theory of change that framed a theory-based evaluation. Most of the evaluation reports claimed to have used a mixed-methods approach. There is a range of usage of 'mixed-methods' in these reports. Majority of the reports followed the GEF guidelines to terminal evaluation mixed with their own institutional guidance where they exist (i.e., for the MIEs), with some mention of the AF guidance and framework.
Dimension 4: Evaluation Methods and Data	All 12 evaluations used desk-based review of project/programme documents and other related literature. Majority of the evaluations combined multiple qualitative data collection techniques such as direct observations, interviews, and group discussions either remotely or in the field.
Dimension 5: Analysis and Findings	The quality of analysis and findings across 10 of the final evaluation reports was generally satisfactory (5 reports were Moderately Satisfactory; 2 reports were Satisfactory; and 3 reports were Highly Satisfactory). These reports presented findings that were balanced, triangulated, clear, and in most parts in-depth. Amongst the techniques used were appreciative inquiry, thematic analysis, descriptive analysis, content analysis, interpretive analysis, and statistical analysis.

Dimension	Summary of findings
Dimension 6: Conclusions, Lessons, Recommendations	9 evaluation reports were satisfactory (3 Moderately Satisfactory and 6 Satisfactory). A common but minor weakness is presenting the lessons as if they were recommendations. The substance of the lessons was generally sound and logically flowed from the findings and conclusion. Common weakness of recommendations includes the need for prioritization and/or clarity on who owns the recommended actions and the need to specify actions that inform future CCA programming.
Dimension 7: Inclusion, Ethics, Independence	10 evaluation reports were rated Moderately Satisfactory (n=4), Satisfactory (n=4), and Highly Satisfactory (n=2) thereby majority of the reports scored well in this dimension. Some evaluations included the sworn statement for ethical evaluation in the annex; others included a sub-section for this purpose. Only two reports were Moderately Unsatisfactory, and the main weakness of these reports is that they did not elaborate on how the evaluation was independent, ethical, and inclusive.
Dimension 8: Management and Governance	8 evaluation reports were in the satisfactory region (Moderately Satisfactory = 5; Satisfactory = 2; and Highly Satisfactory = 1) and only had moderate to no shortcomings. For the 4 unsatisfactory (Moderately Unsatisfactory = 3; and Unsatisfactory = 1) reports, the most pressing issue was the lack of guidance from the IE. One report was particularly descriptive and demonstrates a lack of proper quality assurance.
Dimension 9: Utility	Majority of the reports were either Moderately Unsatisfactory (n=5) or Moderately Satisfactory (n=4); three reports that were Unsatisfactory (n=1) and Satisfactory and Highly Satisfactory (n=2). The worst performing reports in this dimension did not identify the main users of the evaluation and what follow up was to be done to ensure uptake of lessons and recommendations.

Quality of the projects/programmes

It is unclear if the projects/programmes have improved. The proportion of projects/programmes in final evaluations that were reviewed with overall ratings of Moderately Satisfactory to Highly Satisfactory were roughly the same for the two syntheses – around 94 per cent (n=15) for AF-TERG’s first synthesis compared to around 90 per cent (n=10) for this synthesis. A higher proportion of projects/programmes (36 per cent) covered in the second synthesis are Highly Satisfactory, compared to 19 per cent of projects/programmes covered in the first synthesis. Half of the projects/programmes in the first synthesis were Satisfactory.

While it is unclear if project/programme quality has improved, on balance, it can be concluded that the quality has not deteriorated. The lowest overall rating for projects/programmes covered by evaluations in this synthesis is Moderately Unsatisfactory (9 per cent), compared to the first synthesis’ Unsatisfactory (6 per cent).

Programmatic challenges and issues

The most common challenges that were noted to have affected the performance and effectiveness of the projects/programmes are as follows:

Flaws in the project design including overambitious geographic scope and targets, lack of stakeholder analysis, needs analysis, and/or capacity gap analysis, and activities and outputs that do not have a clear link with the intended outcomes. These flaws are closely related to another recurring issue, the **lack of understanding of the local context** including of ecosystem, environment, and communities.

Operational and logistical issues such as poor quality or high turnover of Implementing Entity staff, cumbersome procurement process, and the timing of release of funds. The slow release of funds came at different levels including from the IE and the AF.

Low capacity of implementing partners to successfully implement and manage projects/programmes, which manifested in the lack of adequate qualified personnel, onerous or bad decision-making that led to delays, and poor reporting.

Weak M&E system, which hampered the ability of projects/programmes to learn, adapt, and improve its operations. Among the factors that contributed to this issue include poor overall design, lack of mechanism to

measure improvements, delayed institutionalization of M&E, inadequate uptake and utilization of the framework and tools, and the inability to hire a dedicated M&E officer and undertake regular monitoring activities due to minimal budget provisions for M&E. There was also reportedly an absence of M&E systems that can monitor and evaluate longer-term adaptation results.

Lack of beneficiary participation leading to non-acceptance of new technologies, products, or livelihoods being introduced.

Lack of a gender and youth perspective in planning and consequently in implementation. This meant that some activities such as training and outputs such as livelihoods have not adequately factored in the differing needs of various groups of beneficiaries.

Beyond the challenges within the control of the projects/programmes, there were also **external factors** that negatively affected performance such as Covid-19, natural disasters, and security issues. The pandemic was noted to have caused some delays and/or procurement issues. Both natural disasters and security issues also slowed down the implementation process.

Lessons from the portfolio

Four recurring lessons related to the common challenges/issues have been identified across the 12 evaluation reports. These lessons also appeared as common lessons in the first synthesis and therefore it raises questions on how these lessons have been communicated to the IEs and how the IEs were incentivised to take up of such lessons.

Lesson 1: The importance of understanding local context and environment. Such understanding and the overall knowledge of the ecosystem, political context, stakeholder and partner capacities, and community needs are highly pertinent to appropriate project design with fit-for-purpose components and implementation structure.

Lesson 2: Inclusion is a key determinant of project/programme effectiveness and sustainability. The multi-dimensional needs and role of women, youth, indigenous people, and other groups must be built into the design, reviewed regularly, and mainstreamed into the implementation processes. A lack of focus on inclusion issues undermines the ability of projects/programmes to mark genuine positive impact on the adaptive capacity and resilience of individuals and communities.

Lesson 3: Strong partnerships and capacities are a pillar of successful implementation. Without the buy-in and participation of capable partners on the ground, relevant institutions, and communities, project/programme efficiency suffer.

Lesson 4: Robust M&E systems facilitate reporting, learning, and adaptive management. However, gaps remain in the specification of targets and indicators, utilization, and generally institutionalization/mainstreaming of a plan that broadly serves adaptation programming needs. Apart from adequate budget provisions within the project/programme timeframe, M&E systems that have a longer time horizon are needed to capture adaptation results that only manifest after completion.

Recommendations

This report presents the recommendations for the AF Secretariat. The internal recommendations for AF-TERG along with the recommendations for the design of the subsequent syntheses have been removed from this version.

Programmatic recommendations for the AF Secretariat

Recommendation 1: Strengthen due diligence of requirements that can foster better understanding of local contexts and environment during project development phase, including stakeholder analysis,

beneficiary needs analysis, capacity gap analysis, and gender and other inclusion studies. The review checklist could include the clear link between the findings coming out of these analyses and the overall theory of change, in both diagrammatic and narrative formats, implementation structures, and risk management mechanisms.

Recommendation 2: Check for the adequacy of the M&E budget vis-à-vis M&E plans at entry and require updates on how M&E is working (or not) in progress reports. Check that there is a sound M&E plan as well as an adequate budget that can support the components of the plan during the design stage. Thereafter, a short section in the progress reporting template could be dedicated on assessing how the M&E is operating.

Recommendation 3: Conduct a benchmarking exercise on the scale of M&E budgets in approved proposals in collaboration with AF-TERG. Given the lack of information in the FE reports, it will be important to understand the scale of the M&E budget of AF-funded projects/programmes in various contexts and sectors. Follow-up interviews with the IEs can then be undertaken to complement the benchmarking exercise and test the adequacy of various budgetary range. This can also inform future guidance either from the AF Board Secretariat or AF-TERG on budgeting for the evaluations.

1. Introduction

The Technical Evaluation Reference Group of the Adaptation Fund (AF-TERG) commissioned a short-term consultant to undertake a synthesis of the AF's final evaluations submitted between July 2020 and August 2023. The assignment started on 16 October 2023. It is comprised of three deliverables: (i) a draft evaluation review methodology, (ii) a draft evaluation synthesis report and improvements to the review methodology, and (iii) final report and methodology.

A draft evaluation review methodology was completed and cleared on 13 November 2023 after a period of revision and pilot application. A draft evaluation synthesis report along with the suggested improvements to the review methodology was submitted to AF-TERG on 30 November 2023.

This report, the final deliverable of the consultancy, is the revised and final version of the draft evaluation synthesis report. It incorporates the feedback that AF-TERG provided between December 2023 and January 2024.

The AF Evaluation Policy and relevant guidance documents served as an anchor for the quality reviews and synthesis in this report. It is important to note that the final evaluations covered in this synthesis were not yet subject to the application of the AF Evaluation Policy¹ but were guided by the preceding guidelines, for example, the AF's Guidelines for Project/Programmes Final Evaluations² and the AF's Evaluation Framework³. As such, the quality review that this consultancy undertook were not intended to present the level of compliance of Implementing Entities, but merely to provide an indication of any capacity gaps that might need to be filled in the future.

2. Background to the evaluation synthesis

Final evaluations are a core element of the learning and accountability architecture within the AF. The AF Operational Policies and Guidelines for Parties to Access Resources from the Adaptation Fund⁴ (para. 66) states that:

[a]ll regular size concrete projects and programmes that complete implementation will be subject to terminal evaluation by an independent evaluator selected by the implementing entity. All small size concrete projects and programmes, as well as readiness grant projects, shall be subject to terminal evaluation if deemed appropriate by the Board and shall follow an evaluation process as decided by the Board using templates approved by the Board. Terminal evaluation reports will be submitted to the Board as stipulated in the project agreement.

The AF Evaluation Policy further reaffirmed this:

“[a]ll Fund-supported projects and programmes that complete implementation should conduct a final evaluation to assess project/programme performance and impact to support learning and accountability, and inform future climate change adaptation (CCA) interventions... All Implementing Entities (IEs) are required to commission an independent final evaluation of their projects, submitted to the secretariat and the Designated Authorities (DA) within nine months of project completion.”

¹ See also ToR, p. 2

² Accessed from AF website: [Guidelines for Project/Programme Final Evaluations - Adaptation Fund \(adaptation-fund.org\)](https://www.adaptation-fund.org/wp-content/uploads/2017/08/Amended-OPG_Oct-2022_2.pdf).

³ Accessed from AF website: [Evaluation Framework - Adaptation Fund \(adaptation-fund.org\)](https://www.adaptation-fund.org/wp-content/uploads/2017/08/Amended-OPG_Oct-2022_2.pdf).

⁴ https://www.adaptation-fund.org/wp-content/uploads/2017/08/Amended-OPG_Oct-2022_2.pdf

With these policy requirements, the AF-TERG seeks to generate evaluative insights and knowledge specifically by articulating and utilizing evaluation results for accountability and learning within the AF as programmed in its indicative work programme for fiscal years 2021 to 2023 (FY21 – FY23)⁵.

The AF-TERG undertook an initial evaluation synthesis in 2020 as part of its initial work progress in Financial Year 2020 and in accordance with the technical group's objectives. The initial synthesis presented the first quality review of AF final evaluation reports. That synthesis covered 17 final evaluations of projects/programmes received from 2015 to June 2020 and was, therefore, based on the first cohort of completed projects in the AF's portfolio.⁶

This second synthesis builds on the AF-TERG's first synthesis. It aims to strengthen the AF-TERG's efforts to provide synthesis products that will strengthen learning and accountability within the AF and beyond. It will also inform the formulation and implementation of the AF-TERG's new multi-year work programme, further the work of the AF and its IEs and, above all, provide an important frame of reference for the AF Evaluation Policy.

3. Purpose, objectives, and scope of the FE synthesis

Purpose

The main purpose of the synthesis is to generate lessons learned from project/programme design and implementation that will feed into and enable the AF to improve current and future programming. It also aims to serve accountability purposes. The primary audience of the synthesis are the AF Board Secretariat and the AF Board. The report may also be useful to secondary audiences including but not limited to Implementing Entities, other climate funds and donors operating in the adaptation space, and other organizations and evaluators aiming to undertake similar syntheses.

Objectives

As per the ToR, the synthesis has four main objectives:⁷

1. To report to the Board on the quality of final evaluations and on the overall performance and effectiveness of completed projects and programmes, including in relation to improved adaptive capacity and resilience.
2. To review the quality of final evaluations reports based on compliance with the criteria as laid out in the Evaluation Policy and Guidance of Final Evaluations
3. To assess the quality of the lessons in the final evaluations' recurring findings, lessons learned and evidence on specific thematic areas of interest such as gender/equity, improved adaptive capacity and resilience, or sustainability.
4. Serve an accountability purpose through exploring the extent of management response to recommendations provided in the Final Evaluations and how the management response is fed into future programming.

Scope

The synthesis covers 12 final evaluations submitted to the AF from July 2020 to August 2023. Details of these projects/programmes are presented in Table 2.

⁵ Fiscal Years 2023 – 2024 Update To The Work Programme Of The Adaptation Fund Technical Evaluation Reference Group (AF-TERG) (https://www.adaptation-fund.org/wp-content/uploads/2022/03/AFB.EFC_.29.7-AF-TERG-Work-Programme-FY23-24-Update_final.pdf)

⁶ ToR, p.3

⁷ ToR, p. 3

4. Key synthesis questions

As the ToR did not specify any synthesis questions, the consultant generated high-level questions that will enable the synthesis to meet its purpose and objectives. The questions are related to two broad aspects of this exercise – the quality of the final evaluation reports and the performance of the projects/programmes. A comparative element focusing on the project/programme evaluation ratings has been introduced to draw on AF-TERG’s first synthesis.

1. What is the quality of the final evaluation reports submitted to the AF?
2. What are the main weaknesses of the evaluation reports based on the review methodology’s quality criteria?
3. How well did the projects/programmes perform overall? Were they effective and if they were, were the results likely to be sustained? Are the longer-term results likely to be achieved? Did they perform better/worse than the portfolio of projects/programmes assessed in AF-TERG’s first synthesis?
4. What are the common challenges (if any) affecting the performance across the portfolio of projects/programmes?
5. What are the common lessons that can be learned from the portfolio of projects/programmes?

5. How to read this report

The succeeding sections of the report will firstly describe the approach that this synthesis adopted (Section 6 Methodology) to provide guidance to the readers as to how the findings and analyses were generated. Section 7 gives a brief overview of the portfolio of evaluations included in the synthesis. Sections 8 to 10 were designed to directly respond to the four objectives of the synthesis as specified in the ToR (Table 1). Finally, Section 11 puts forward recommendations for future programming as well as for the improvement of the review methodology.

Table 1 Mapping of report sections against the synthesis objectives

Report section	Synthesis objectives
Section 8 Quality of the Final Evaluation Reports	Objectives 1, 2, 4
Section 9 Quality of the Projects/Programmes	Objective 1
Section 10 Lessons from the Synthesis	Objective 3

6. Methodology

The methodology used for this synthesis does not follow one single design but draws on the combined strengths of realist synthesis⁸ and lessons learned synthesis⁹. This complementation has allowed the synthesis to be pragmatic and flexible in the face of constraints and limitations (see Limitations) to meet the objectives of the assignment.

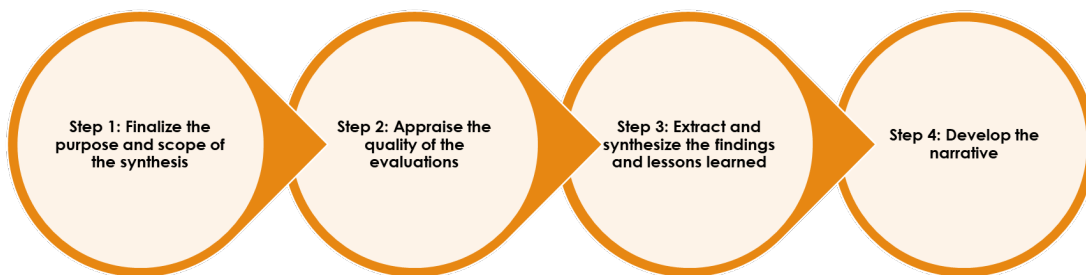
⁸ Realist synthesis focuses on bringing together evidence to come up with policy-based recommendations. It allows for thematic grouping of results and practical recommendations (FAO Evaluation Synthesis Guidance, p. 3).

⁹ ‘Lessons learned synthesis is a product of the evaluative process of reflecting experiences in undertaking an evaluation. Lessons can take the form of describing what should or should not be done or describing the outcome of different processes’ (FAO Evaluation Synthesis Guidance, p. 4).

It is worth noting that ‘there is not one specific definition for various types of evaluation synthesis. Organizations typically tailor them to their own needs and policy’.¹⁰ The critical feature of an evaluation synthesis is it captures ‘evaluative knowledge and lessons learned on a certain topic from a variety of existing evaluations through aggregated and distilled evidence in order to draw more informed conclusions (and sometimes recommendations) on a specific topic or question’.¹¹

This pragmatic methodology was applied to the synthesis in four steps, which is a minor departure from the suggested three-step methodology in the ToR. The change was suggested and made to strengthen the lessons learning function of the synthesis. The steps are elaborated below and as illustrated in Figure 1.

Figure 1 Four-step methodology of the synthesis



Step 1: Finalize the purpose and scope of the synthesis.

At the beginning of the assignment, the consultant agreed with AF-TERG on the purpose, scope, and primary/secondary audience of the synthesis. The main purpose of the synthesis was also consulted with the AF Board Secretariat and received no objections.

The ToR only requires a methodology for reviewing the quality of the final evaluations, but not for the synthesis. As such, it was important to agree that a more holistic synthesis methodology that covers both quality review and synthesis is necessary. There was also a need to define more precisely the primary and secondary users of the report to produce a more focused output.

Step 2: Appraise the quality of the final evaluations.

The synthesis literature systematically includes the assessment of the evidence base that will be used for the synthesis. The FAO noted that syntheses should include an ‘evaluation of evaluations’, adding that well-defined inclusion/exclusion quality criteria is key.¹² Some organizations¹³ adopt a formal framework that allows for the assessment and rating of quality, and which specifies a threshold for inclusion in the synthesis.

This synthesis follows this good practice of setting out a transparent quality review framework (please refer to Annex 2). Hence, prior to assessing the evaluations, the first activity under this step is to develop a quality review framework. For this assignment, the consultant decided against an inclusion/exclusion threshold. This is a pragmatic decision to ensure that lessons can be learned even from weaker evaluation reports but noting that only findings and lessons that were adequately substantiated could be included in the analysis.

¹⁰ FAO Evaluation Synthesis Guidance, p. 3

¹¹ The Presidency of the Republic of South Africa, 2014 in FAO Evaluation Synthesis Guidance, p. 3

¹² FAO Evaluation Synthesis Guidance, p. 8

¹³ WFP, IFAD, and FAO for example

- **Development of a quality review framework**

The ToR states that the review methodology will ‘draw upon existing approaches’. The methodology shall be fully aligned with the AF Evaluation Policy, its relevant guidance notes, and other relevant documents of the Fund. The consultant reviewed the new AF Evaluation Policy along with the guidance¹⁴ documents, the AF’s Evaluation Framework and Guidelines for Project/Programme Evaluations, and the 2021 Synthesis of Final Evaluations report. To learn from how other organizations are undertaking this type of quality assurance or assessment of evaluation reports, the consultant also considered existing approaches and frameworks where they were publicly available such as those from the Global Environment Facility, Green Climate Fund, World Fund Programme, United Nations Population Fund, and the UK Foreign, Commonwealth & Development Office.¹⁵ The process of developing the draft methodology is set out in detail in Annex 2.

Following the review, the consultant distilled the AF Evaluation Policy’s requirements and their operationalization in the relevant guidance documents. This exercise entailed translating the AF Evaluation Principles into quality dimensions that form the organizing structure of the proposed methodology. Under each principle, the requirements of the AF Evaluation Policy and their operationalization in the guidance notes were generated.

AF Evaluation Policy’s 7 Evaluation Principles: (i) Relevance and utility; (ii) Credibility and robustness; (iii) Transparency; (iv) Impartiality and objectivity; (v) Equitable and gender-sensitivity inclusivity; (vi) Complementarity and (vii) Complexity-sensitive and adaptive.

Nine quality dimensions were selected to represent comprehensively the requirements of the AF Evaluation Policy. These dimensions were then broken down into more specific considerations for transparency in the assessment process.

1. Structure and clarity
2. Context, purpose, scope, and objectives
3. Evaluation framework and methodology
4. Methods and data
5. Analysis and findings
6. Conclusions, lessons, and recommendations
7. Inclusion, ethics, and independence
8. Management and governance
9. Utility

¹⁴ Such as the Guidance supporting the operationalization of the (i) Evaluation Principles; (ii) Evaluation Criteria; (iii) Evaluation Budgeting; (iv) Evaluation Reporting; and (v) Final Evaluations.

¹⁵ Note that most of these existing approaches are applied to individual evaluation reports upon submission as part of a quality assurance process, as opposed to a group of evaluations for quality assessment and synthesis purposes as in this assignment. There is also a difference on who performs the reviews. For example, the GCF Secretariat assess all mid-term and final evaluations through internal staff resources. The UK FCDO and WFP outsource the function to a firm or a consortium of firms.

Finally, the methodology adopted an even numbered Likert scale with six rating levels.¹⁶ This follows the AF's 'mandatory rubrics rating scale' for final evaluations required to assess and communicate the performance of a project/programme against the Evaluation Policy's evaluation criteria.¹⁷

Each of the quality dimension ('criterion') was rated 1-6 depending on the extent of the evaluation reports' shortcomings in meeting the criterion. The overall score is calculated using weights. The weights reflect the level of importance of each criterion in the assessment.

The biggest importance was accorded to Methods and data, Analysis and findings, and Conclusions, lessons, and recommendations (20 points each). The robustness of methods and data, and how they were used in the analysis and to inform the lessons and recommendations are highly pertinent to an evaluation's usefulness and credibility; they are the practical application of the Methodology and framework and Inclusion, ethics, and independence (10 points each). The other dimensions were assigned five points each.

The full set of quality dimensions and specific recommendations, the rating scale, and the weights together make up the accompanying tool in Excel spreadsheet format. This tool was developed to facilitate the assessment in the most transparent manner and is consistent with the current practices of other organizations. Relevant formulae were embedded in the spreadsheet to automate calculations for reviewers.

- **Assessment of the final evaluations' quality**

The quality of the 12 final evaluations was assessed using the methodology previously described. A pilot was first undertaken to inform the fine-tuning of the review methodology. Four evaluation reports were subject to the pilot undertaken by two reviewers. A moderation session helped to compare scores and notes on the improvements required to the review methodology. Following the pilot and revisions to the methodology, the remaining evaluation reports were assessed, and the findings analysed using descriptive and summative statistics.

Step 3: Extract and synthesize findings and lessons learned and Step 4: Develop the narrative.

The final two steps to the synthesis include the analysis of findings, generation of common lessons and recommendations, and the writing of the report. The analysis was guided by the synthesis questions. It adopted a qualitative process.

To analyse the quality of the projects/programmes, the synthesis considered the ratings for each evaluation criterion in the individual reports. They were extracted and recorded in Excel. Where there were differing metrics used, for example a mix of 4-point and 6-point scales, the scales were equalised to a 4-point scale based on the evidence presented in the reports. Note that this only applied to Relevance, Sustainability, and Impact.

Of note is that out of the 17 evaluations covered by the first synthesis, there was no available report for the Cook Islands project and hence 16 out of 17 project ratings were extracted. And out of the 12 evaluations for the second synthesis, the project in Rwanda did not have clearly outlined ratings and hence was excluded.

¹⁶ The original proposal of the consultant to adopt four-rating levels was revised to six after a pilot as the latter captures more gradient in the quality of the evaluation reports. The four-rating scale was proved to be either too punitive or too generous.

¹⁷ Guidance in Support of the Operationalization of the Evaluation Policy: Final Evaluations, p. 15

Once all the ratings for all the indicators were added to Excel, calculations were done to assess the frequency of the various ratings within the 16 reports in the first synthesis and the 11 reports in this synthesis respectively. A comparison of percentages made it possible to visualize and compare the quality of projects/programmes between the two syntheses.

Challenges and lessons learned were organized along thematic lines where they were cited across evaluation reports. Findings and lessons that were mentioned by half of the reports (n=6) or more were all included in the reporting process.

A rapid thematic coding technique was used to generate these common challenges and lessons. The 12 reports were merged into one PDF file. Each evaluation was scanned to identify common thematic areas that challenged the projects/programmes at all stages of planning and implementation. As the themes emerged from each of the 12 reports, they were combined under colour codes and noted in a separate Excel sheet. On completion of the readings, the group of colours for the challenges were assigned title and 11 of the most important/ common challenges were represented in a frequency table.

Similarly, for the common lessons, themes were also generated and added under specific colours in Excel and later combined in a frequency table.

Limitations

The synthesis solely relied on secondary information in the 12 evaluation reports. The ToR did not have provisions for follow-up interviews with Implementing Entities or evaluators and therefore there was no opportunity for verification and triangulation as well as further consultations to enrich the findings and recommendations.

Performance assessments of both the evaluation reports and the projects/programmes were limited to ratings comparison. The two syntheses adopted a different review methodology to assess quality of final evaluation reports and thus they are not comparable. The project/programme ratings were compared, but the limited resources did not allow for a deeper examination. Further, this synthesis also adopted a different quality assessment methodology from the AF-TERG's Rapid Evaluation of the AF's evidence gap mapping.¹⁸

While the number of evaluation reports covered by this synthesis could be perceived to have a negative impact on the findings and their usefulness, there is no minimum number of reports in the literature for synthesis exercises. It is not unusual in research syntheses for search parameters to yield thousands of literatures then end up with five sources for inclusion. It is often more about the quality of the evidence base, than the sheer number of reports.

7. Portfolio overview

This section presents a brief overview of the portfolio of projects/programmes that were covered by the evaluation reports included in this synthesis.

The evaluation synthesis included the final evaluation reports of 12 completed projects/programmes, as presented in Table 2. Two of the projects/programmes were implemented in Asia (Cambodia, Uzbekistan), three in Latin America (Colombia, Cuba, Uruguay), and seven in Sub-Saharan Africa (Egypt, Ghana, Madagascar, Mali, Mauritius, Rwanda, South Africa).

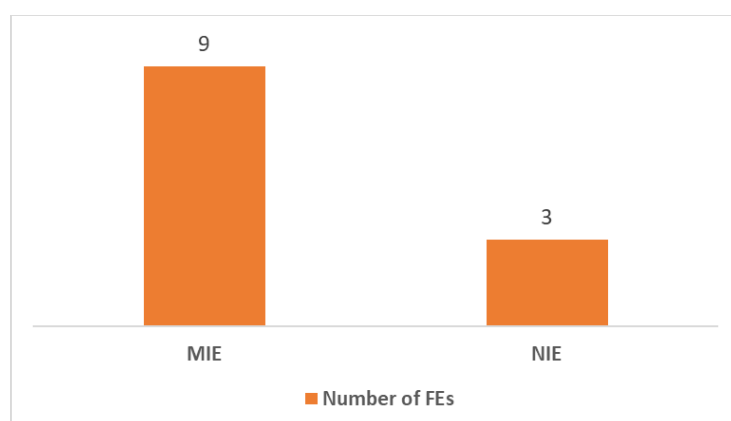
¹⁸ Rapid Evaluation of the Adaptation Fund (18 August 2023 draft), accessed from [AFB.EFC .32.6.Rev .1 Rapid-evaluation.pdf \(adaptation-fund.org\)](https://www.afb.efd.org/32.6.Rev.1.Rapid-evaluation.pdf)

Table 2 Projects/programmes included in the synthesis

	Project name	Country
1	Enhancing Climate Change Resilience of Rural Communities Living in Protected Areas of Cambodia	Cambodia
2	Risk and vulnerability reduction towards climate change in the Momposina Depression region in Colombia	Colombia
3	Reducing vulnerability to coastal flooding through ecosystem-based adaptation in the South of Artemisa and Mayabeque provinces of Cuba	Cuba
4	Building Resilient Food Security Systems to Benefit the Southern Egypt	Egypt
	Increased Resilience to Climate Change in Northern Ghana through the Management of Water Resources and Diversification of Livelihoods	Ghana
6	Promoting Climate Resilience in the Rice Sector through Pilot Investments in Alaotra-Mangoro Region	Madagascar
7	Programme Support for Climate Change Adaptation in the vulnerable regions of Mopti and Timbuktu (PACV-MT)	Mali
8	Terminal Evaluation of “Climate Change Adaptation Programme in the Coastal Zone of Mauritius”	Mauritius
9	Reducing Vulnerability to Climate Change in North West Rwanda through Community Based Adaptation	Rwanda
10	Taking Adaptation to the Ground: A Small Grants Facility for Enabling Local Level Responses to Climate change	South Africa
11	Building resilience to climate change and variability in vulnerable smallholders	Uruguay
12	Developing climate resilience of farming communities in the drought prone parts of Uzbekistan	Uzbekistan

Of the 12 projects/programmes, nine were implemented by Multilateral Implementing Entities (MIEs) (six UNDP, two UNEP, and one WFP), and three were implemented by National Implementing Entities (NIEs) (South African National Biodiversity Institute (SANBI), Ministry of Energy (MoE) Rwanda, Agencia Nacional de Investigacion e Innovacion (Uruguay)).

Figure 2 Implementing entities

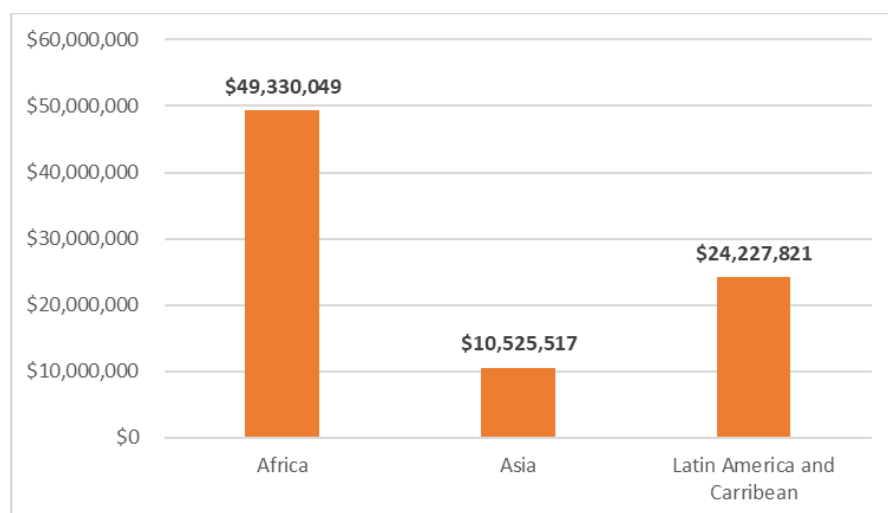


The AF’s grant to these projects/programme range from US\$ 2,442,682 (South Africa) to \$10,000,000 (Rwanda) (Table 3). Regionally, the AF grant to the projects/programmes in Africa is a total of \$49,330,049, followed by Latin America and the Caribbean with the sum of \$24,227,881, and lastly by Asia with \$10,525,517 (see Figure 3).

Table 3 AF Grant received by the projects/programmes evaluated by FEs within the scope of this synthesis, by country

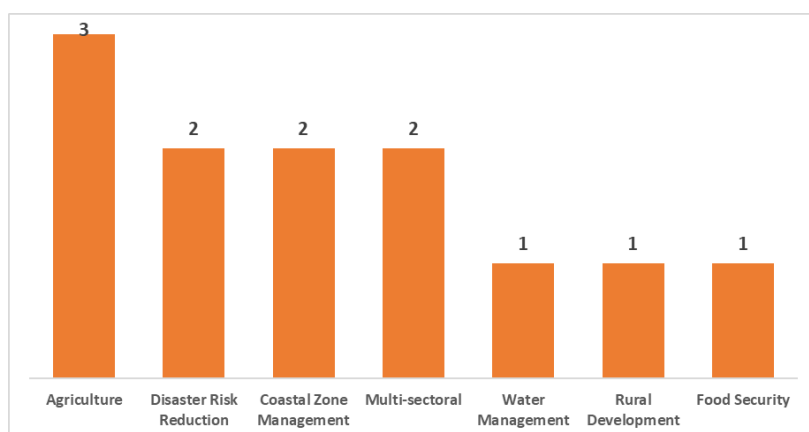
Country	AF Grant
South Africa	\$2,442,682
Madagascar	\$4,705,000
Cambodia	\$4,954,273
Uzbekistan	\$5,571,244
Cuba	\$6,067,320
Egypt	\$6,904,318
Mali	\$7,864,837
Ghana	\$8,293,972
Colombia	\$8,500,000
Mauritius	\$9,119,240
Uruguay	\$9,660,501
Rwanda	\$10,000,000

Figure 3 Aggregate AF Grant to projects/programmes evaluated by FEs within the scope of this synthesis, by region



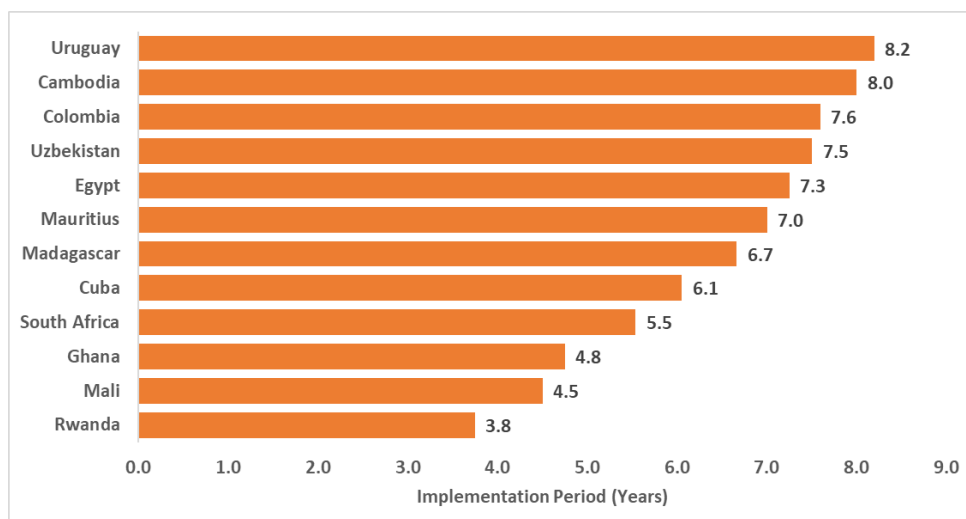
These projects/programmes were spread across several sectors including agriculture, disaster risk reduction, coastal management, and others as per the AF classification.

Figure 4 Sector classification of the projects/programmes in the synthesis portfolio



These projects were implemented over a period of about four years to over eight years. As to be discussed in the analysis, some projects/programmes experienced delays for a variety of reasons. These reasons are a mix of internal and external factors (see Figure 29).

Figure 5 Implementation period of the projects/programmes



Four countries (Cambodia, Madagascar, Mali, South Africa) experienced substantial start-up delays ranging from 140 days to over 240 days (Table 4).

Table 4 Project/programme start-up delays

Country	Planned Start Date	Actual Start Date	Delays (#days)
Cambodia	01/01/2013	21/05/2013	140
Colombia	28/06/2012	28/06/2012	0
Cuba	30/09/2014	01/10/2014	1
Egypt	01/03/2013	01/03/2013	0
Ghana	01/04/2016	01/04/2016	0
Madagascar	01/02/2012	01/10/2012	243
Mali	01/10/2015	04/03/2016	155
Mauritius	30/08/2012	30/08/2012	0

Rwanda	01/06/2014	01/06/2014	0
South Africa	01/04/2015	16/09/2015	168
Uruguay	21/10/2012	21/10/2012	0
Uzbekistan	01/05/2014	01/05/2014	0

Meanwhile, all the projects/programmes had to be extended by a minimum of 274 days (Ghana) to a maximum of 1,280 days (Uruguay) (Table 5). As Figure 6 demonstrates, there is a big divergence between the planned and actual closing dates.

Figure 6 Planned versus actual closing date of projects/programmes

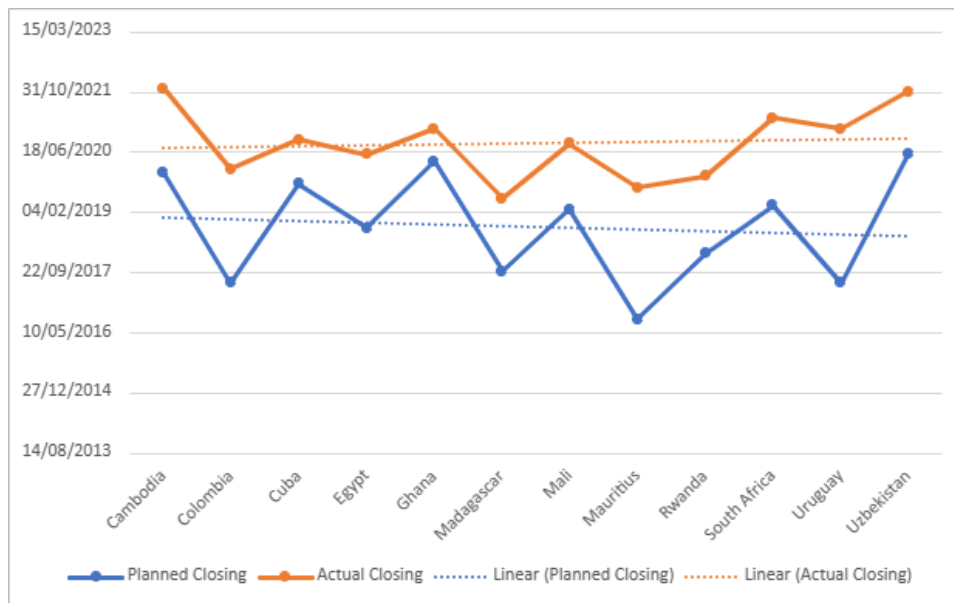
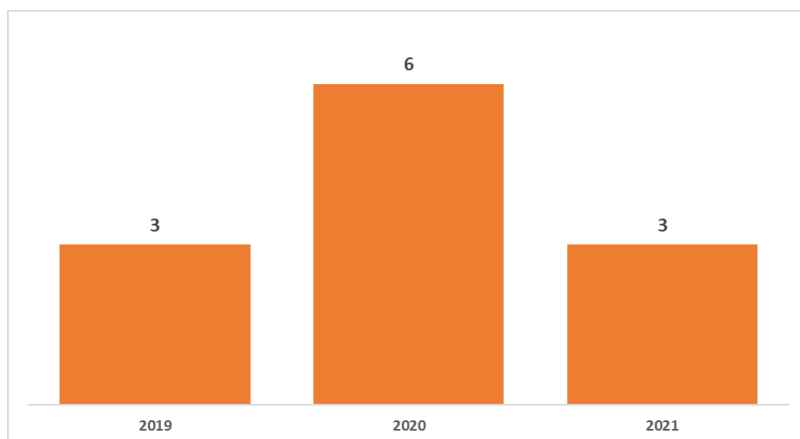


Table 5 Magnitude of extension in the implementation period of the projects/programmes

Country	Planned Closing	Actual Closing	Extension
Cambodia	31/12/2019	01/12/2021	701
Colombia	27/06/2017	31/01/2020	948
Cuba	30/09/2019	30/09/2020	366
Egypt	01/10/2018	01/06/2020	609
Ghana	01/04/2020	31/12/2020	274
Madagascar	01/10/2017	01/06/2019	608
Mali	01/03/2019	01/09/2020	550
Mauritius	31/08/2016	31/08/2019	1095
Rwanda	01/03/2018	02/12/2019	641
South Africa	01/04/2019	31/03/2021	730
Uruguay	30/06/2017	31/12/2020	1280
Uzbekistan	31/05/2020	01/11/2021	519

Fifty per cent of these projects/programmes were completed in 2020, and the rest are equally divided between completion in 2019 and 2021.

Figure 7 Year of completion



8. Quality of final evaluation reports

This section responds to synthesis questions 1 and 2.

1. What is the quality of the final evaluation reports submitted to the AF?
2. What are the strengths and weaknesses of the evaluation reports based on the review methodology's quality criteria?

It fully addresses the ToR's Objectives 2 and 4, and partly addresses Objective 1 (i.e., to report to the Board on the quality of the final evaluations) and Objective 3 (i.e., assessing the quality of lessons learned).

Objective 2: To review the quality of final evaluations reports based on compliance with the criteria as laid out in the Evaluation Policy and Guidance of Final Evaluations.

Objective 4: Serve an accountability purpose through exploring the extent of management response to recommendations provided in the Final Evaluations and how the management response is fed into future programming.

In the assessment, the rating system in Table 6 was applied to each quality criterion. In this section, the individual as well as overall ratings from the quality review (Step 2 of the Methodology) are presented.

Table 6 Rating system for each quality criterion

Rating	Description	Explanation
6	Highly Satisfactory	The criteria were fully met or exceeded and there were no shortcomings.
5	Satisfactory	The criteria were met with only minor shortcomings.
4	Moderately Satisfactory	The criteria were partially met with moderate shortcomings.
3	Moderately Unsatisfactory	The criteria were partially met with noticeable shortcomings.
2	Unsatisfactory	The criteria were somewhat met with major shortcomings.
1	Highly Unsatisfactory	There were severe shortcomings in meeting expected standards.

A note on the quality. Majority (n=8) of the evaluation reports covered in this synthesis were submitted in 2020. This evaluation completion time frame meant that around 42 per cent (n=5) of the evaluations were undertaken during the pandemic. Such a significant external factor must be borne in mind in the reading of the overall quality of the reports. The five evaluations had to adapt their approach to the context and relied on remote data collection coupled with limited field visits undertaken mainly by a local consultant where possible.

Figure 8 Evaluation completion year

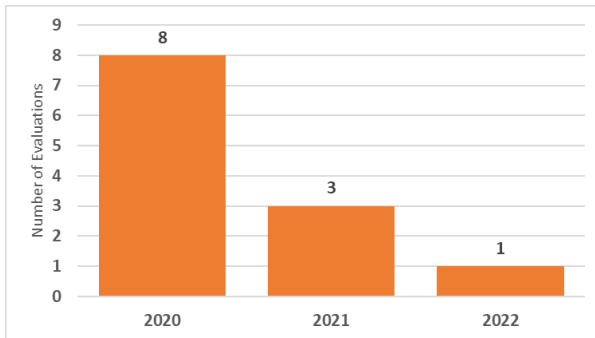
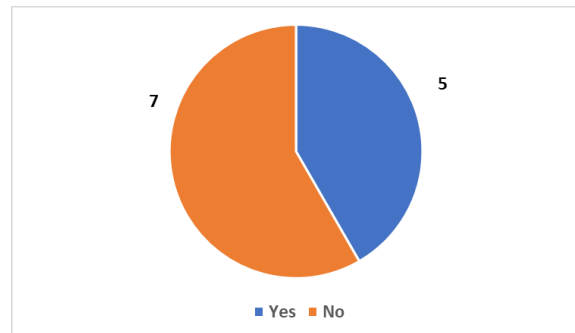


Figure 9 Evaluations undertaken during the pandemic



Dimension 1: Structure and clarity of the report

Dimension 1 assesses the logical structure, accessibility, and comprehensiveness of the final evaluation reports. It looks at the coherence and flow of the content between and within sections, the length of the executive summary and the entire report, and the annexes that have been included to support the report.

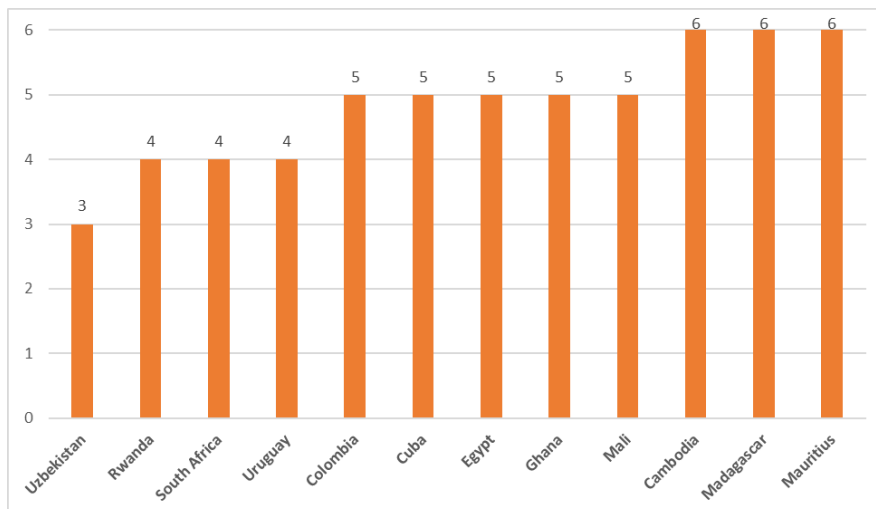
The evaluation reports (n=11) demonstrated moderate to no shortcomings in this dimension hence scoring between Moderately Satisfactory to Highly Satisfactory. The MIEs typically include a sample outline in the evaluation ToR, which consultants followed in the reporting. The other evaluation reports did not deviate from the typical structure of an evaluation report that starts with background information including context, scope, and objective, and followed by methodology, analysis, lessons, and recommendations. The reports are generally accessible to the readers and the narratives are easy to follow.

Example of a Highly Satisfactory Report in Dimension 1 Structure and Clarity of the Report

The structure of the report is logical and complies with the required content of the IE. It was very well written despite being long at over 120 pages for the main text alone. Despite the length, the reviewer found the report very easy to follow as the writing style is accessible even when describing technical, rice farming-specific details. The length could be further justified by the comprehensive substantiation of ratings, not all of which were positive and therefore warranted proper explanation. There was excellent cross-referencing across sections to avoid repetitions, and to annexes proving that every annex was useful to support some claim or other parts of the main text. The Executive Summary was brief and to the point at 4 pages. The consultant also included a French translation for accessibility to the locals.

The main weaknesses of the lone evaluation report that scored Moderately Unsatisfactory in this dimension includes an executive summary that is as detailed as the main findings, the use of jargon (e.g., on conservation agriculture) without explanation in the text, making the report difficult to read, and the excessive use of annexes (e.g., an annex particularly on 'extra information' was included but it is not clear what purpose it serves).

Figure 10 Ratings for Dimension 1 Structure and Clarity



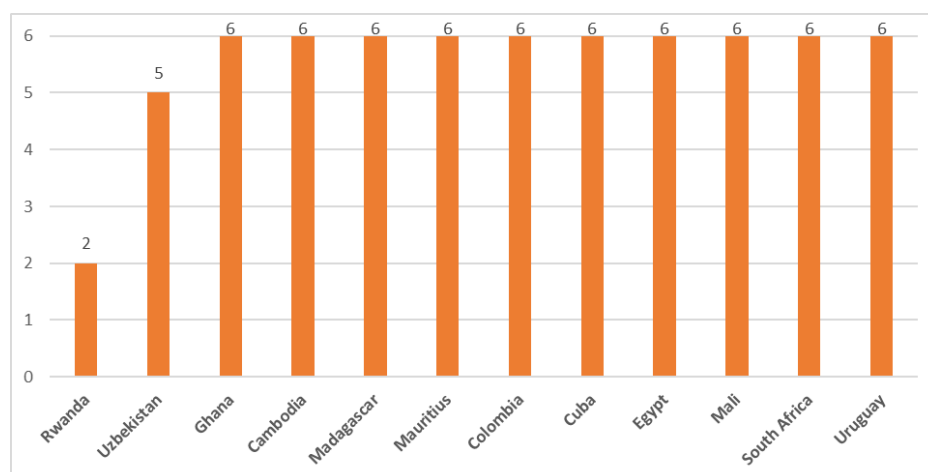
Dimension 2: Context, purpose, scope, and objective

Dimension 2 assesses the extent to which an evaluation report has demonstrated sufficient understanding of the project/programme and its broader linkages with specific organizations, sectors, thematic areas, or geographic space. It also investigates how clearly the purpose, scope and limitations, and objectives including who the primary and secondary audience of the evaluation have been set out.

Majority (n=10) of the evaluation reports scored Highly Satisfactory in this dimension. These reports typically situate the problem the project/programme is trying to address within a context that has been described comprehensively. This is not a surprising finding because normally the contextual material is laid out in project documents whereas the purpose, scope, and objective are outlined in the evaluation ToR.

One report was rated Unsatisfactory in this dimension. The issue with the report is not whether these elements have been included or not, but on how the information in the relevant section was presented. They could have been more accurate or tailored more to the implementation context. For instance, the report covered the geographic and thematic scope, but it stated that the evaluation scope is the whole country when the project under evaluation was implemented in the northwest. The project description in the Introduction section was also a mix of project information, strategies, findings, and activities, which was separate from a project description and development context sections.

Figure 11 Ratings for Dimension 2 Context, purpose, scope, and objective



Dimension 3: Evaluation framework and methodology

Dimension 3 assesses if the evaluation framework and methodology are specific and tailored to the evaluation context, and sufficient to address the purpose, scope, and objectives of the evaluation. The assessment considers the AF evaluation criteria used, as set out in the AF Evaluation Policy and the evaluation questions and where applicable sub-questions to evaluation of a project/programme against the selected evaluation criteria. It also reviews the chosen methodology to address the evaluation questions, whether they provide for multiple lines of enquiry to enable verification and triangulation of results and demonstrate understanding of the methodological limitations.

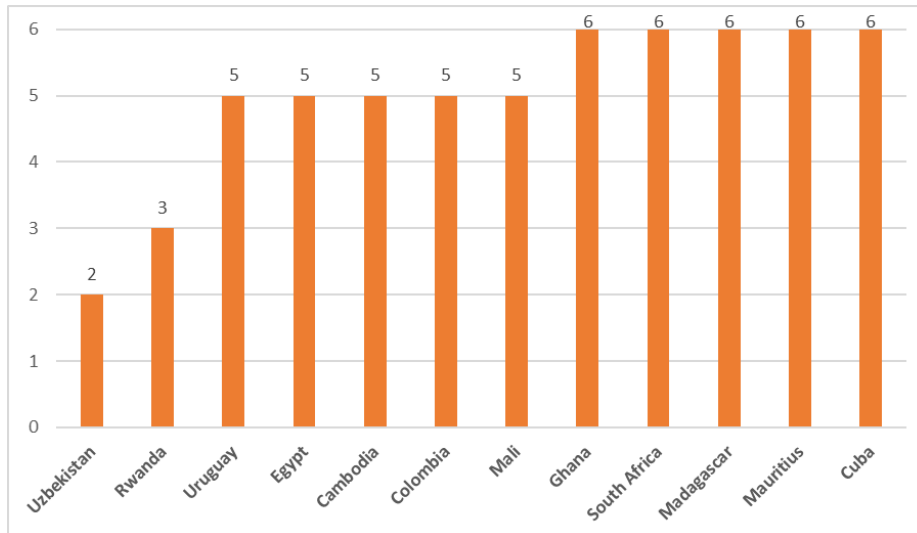
An important caveat to this dimension is that while the assessment checked for the use of the AF Evaluation Policy evaluation criteria, the scoring did not penalise reports when they did not apply such criteria and secured clearance from AF-TERG. This requirement is new, and the policy did not apply to the reports assessed. Not applying the criteria also does not necessarily equate to a lack of capacity of the IEs to apply them. It is most likely only indicative of the fact that there was no formal policy that mandated the IEs to use them. Majority of the reports followed the GEF guidelines to terminal evaluation mixed with their own institutional guidance where they exist (i.e., for the MIEs). Some reports also explicitly mentioned alignment with the AF's policies and guidelines such as "Guide for final evaluations of programs or projects" (e.g., Colombia, South Africa, Egypt) and the AF Evaluation Framework (e.g., Uruguay). Hence, the evaluation criteria applied in these reports followed these earlier guidance documents. MoE Rwanda used OECD-DAC evaluation criteria only.

For most of the evaluation reports (n=11), the evaluation methodology and framework were clearly specified and described. They included either an evaluation matrix (n=8) in the annex or a reconstructed theory of change that framed a theory-based evaluation. The evaluation matrix helped in transparently presenting the judgement criteria or indicators, data sources, and data collection methods used for each of the evaluation questions. One evaluation report that was rated Moderately Unsatisfactory departed from this – it did not specify how success or failure in each criterion was judged. It also claimed triangulation to increase validity and reliability of findings but did not describe how exactly this was done; it also was not evident in the report.

In the case of the only Unsatisfactory report in this dimension, the evaluation matrix only contained the questions and there was no information on the data collection and analytical methods as well as sources for each question. The report noted that it utilized a 'multi-level mixed evaluation' but how

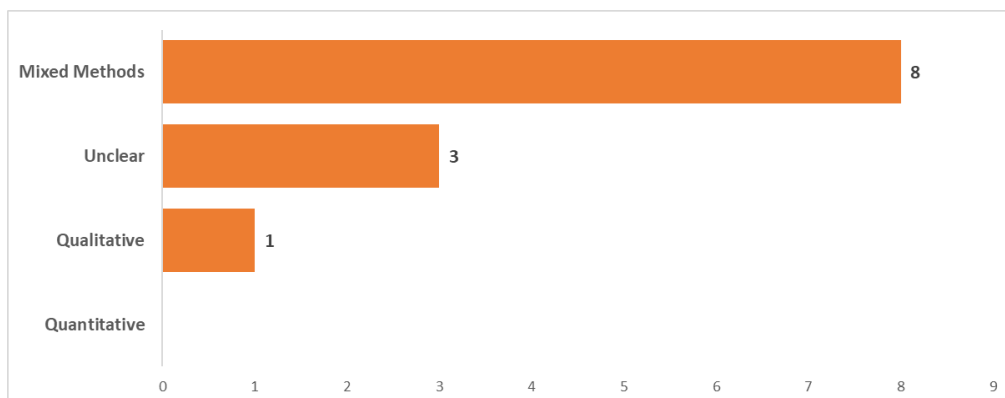
this was applied in the evaluation was neither explained nor was it demonstrated in the report. The methodological limitations were limited to Covid restrictions, which prevented the team leader from traveling. It was not clear though whether it influenced the quality of data and information they managed to collect, as this was not explained in the report.

Figure 12 Ratings for Dimension 3 Evaluation framework and methodology



Most (n=8) of the evaluation reports claimed to have used a mixed methods approach. There is a range of usage of ‘mixed methods’ in these reports. Typically, a mixed methods design combines qualitative and quantitative data collection methods and analytical techniques in a single study.¹⁹ Only three (Cambodia, Ghana, and Uruguay) of the 12 evaluations undertook a survey to complement key informant interviews, focus group discussions, and/or direct observations. The other reports made use of limited secondary quantitative data such as financial information and monitoring data. Three reports (Colombia, South Africa, and Uruguay) did not explicitly specify the methodology used in the report hence they were classified as ‘unclear’ in Figure 13. Upon closer examination on the precise methods used, they were either qualitative in the case of Colombia and South Africa, and mixed method in the case of Uruguay.

Figure 13 Methodology adopted

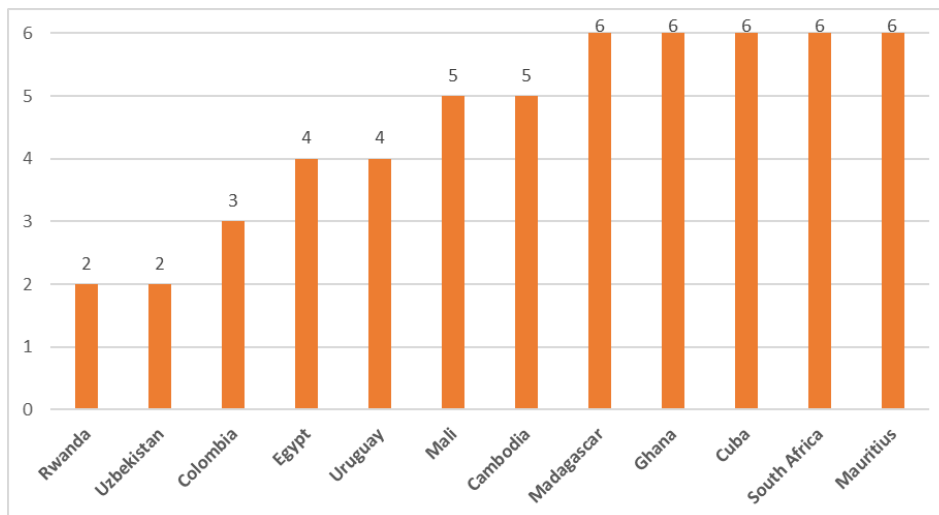


¹⁹ See for example Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Dimension 4: Evaluation methods and data

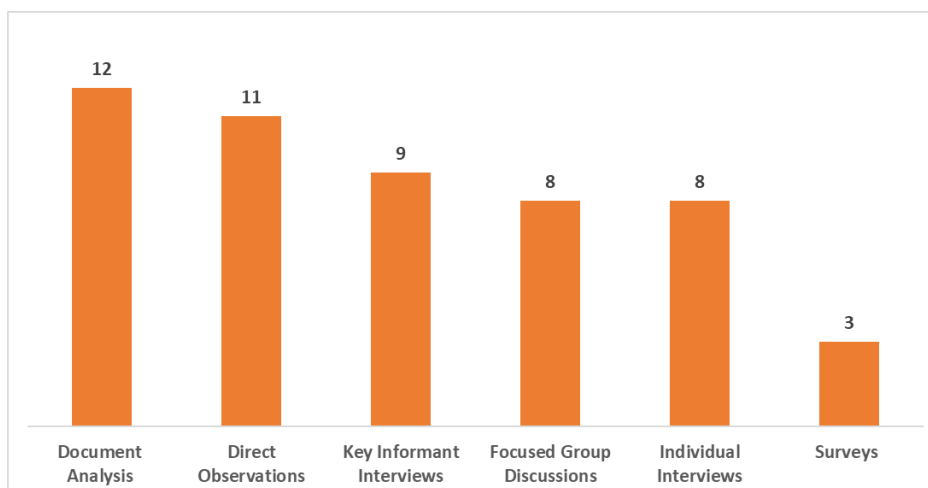
Dimension 4 assesses the appropriateness of the selected methods to adequately address the evaluation questions. It ensures that data sources are adequate and robust, which means that there were appropriate sampling methods where needed, primary and secondary data were used, stakeholders at all levels were reached and consulted, and the methods allowed for the collection and analysis of disaggregated data to show differences between groups where applicable.

Figure 14 Ratings for Dimension 4 Evaluation methods and data



All the 12 evaluations used desk-based review of project/programme documents and other related literature. Majority of the evaluations combined multiple qualitative data collection techniques such as direct observations, interviews, and group discussions either remotely or in the field. Note that in eight of the evaluations, individual interviews were used separately from key informant interviews and implicitly denotes the individual interviews with beneficiaries.

Figure 15 Methods used in the evaluations



The evaluation reports did not explicitly specify the sampling techniques. Based on the narratives, it can be deduced that they all adopted purposive sampling as stakeholder selection was informed by the IEs at first instance. The key informant interviews ranged from 25 to 43 individuals. One evaluation reached 305 beneficiaries as part of a survey. While there could be some improvements to the number of stakeholders reached, for instance to 25 individuals, the evaluation reports typically

offer a justification for the sample size as well as the strategic positioning of these stakeholders to inform the evaluation.

Dimension 5: Analysis and findings

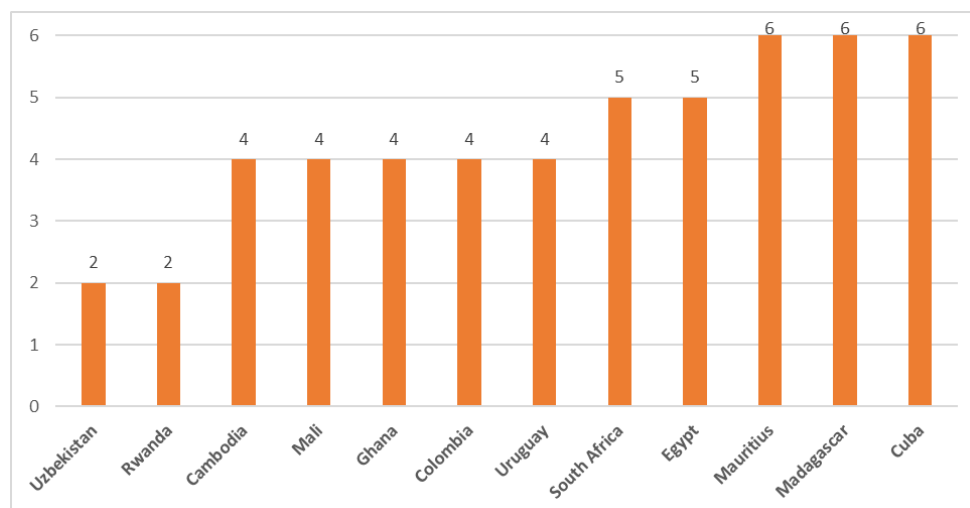
Dimension 5 ensures that the findings and final ratings were well substantiated and that these findings were balanced and supported by legitimate data sources. In considering this quality dimension, the review investigated how clearly the report used the judgement criteria for assessing the evaluation questions and evaluation criteria, and the extent to which the rubrics' rating scale was applied properly. The review also considered how the evaluation analysed potential unanticipated or negative consequences.

The quality of analysis and findings across 10 of the final evaluation reports was generally satisfactory (five reports were Moderately Satisfactory; two reports were Satisfactory; and three reports were Highly Satisfactory). These reports presented findings that were balanced, triangulated, clear, and in most parts in-depth. Fifty per cent (n=6) of the reports did not specify the analytical techniques applied in the evaluation. Among the techniques used were appreciative inquiry, thematic analysis, descriptive analysis, content analysis, interpretive analysis, and statistical analysis.

Example of a Highly Satisfactory Report in Dimension 5 Analysis and Findings

The application of the methods was systematic and transparent, if basing it on the evaluation matrix. There was no survey involved and therefore the evaluation does not necessarily require a sampling technique. The selection of stakeholders was informed by UNDP and the evaluation managed to reach stakeholders at various levels from the government ministries to beneficiaries on the ground. There were not much disaggregated results reported but it was strong in bringing out gender components of both the results and processes.

Figure 16 Ratings for Dimension 5 Analysis and findings



Only two reports were Unsatisfactory in this dimension. The key issues in these evaluations include findings that were incomplete and partially substantiated as well as inaccurate use of evaluation criteria.

For instance, in one of the reports, there was little evidence in the reported results, i.e., of a significant shift in means of livelihood away from agriculture, but it did not delve into why this was the case and what impact it had to the overall achievement of the project. The report also made

claims that raised more questions than answers. For example, the report stated that there were challenges to create diversified means of livelihoods in the new settlements supported by the project – and remarked that ‘this was always going to be a challenge but will likely change for the better as increased economic activity becomes manifest’.

Another weakness is the systemic presence of internal inconsistency. For example, the rating given for the quality of the project design in one of the reports was Moderately Satisfactory and the weaknesses provided included the lack of full stakeholder analysis, capacity assessment of the project partners, and of project and environmental and social risks. However, in the Effectiveness assessment, it turned out that the reason some Outcomes were only partially achieved was because the activities supported were not strongly linked to forest conservation either directly or indirectly, which is such an important design flaw. The report also alluded to the possibility that the low survival rate of replanted trees was because the selected plant species were incompatible with the soil quality, which again is a design flaw.

The two unsatisfactory reports also did not demonstrate any triangulation of findings against various sources.

Dimension 6: Conclusions, lessons, recommendations

Dimension 6 rates how logical, coherent, well-substantiated, and practical/realistic the conclusions, lessons, and recommendations are. It also considers the specification of forward-looking lessons that can help improve and instruct future climate change adaptation programming, and specifically that speak to issues related to maladaptation, CCA finance, and other key issues in CCA. The recommendations were expected to be specific, realistic for implementation and credible given time, resources, and capacities for implementation.

While the ratings across the 12 evaluation reports are still largely (nine of 12) Moderately Satisfactory (n=3) and Satisfactory (n=6), there is not a single report that did not have any shortcoming in this dimension. A common but minor weakness is presenting the lessons as if they were recommendations. This weakness can be easily improved by revising how they were phrased. The substance of the lessons was found to be generally sound and logically flowed from the findings and conclusion. On balance, **the quality of lessons is satisfactory.**

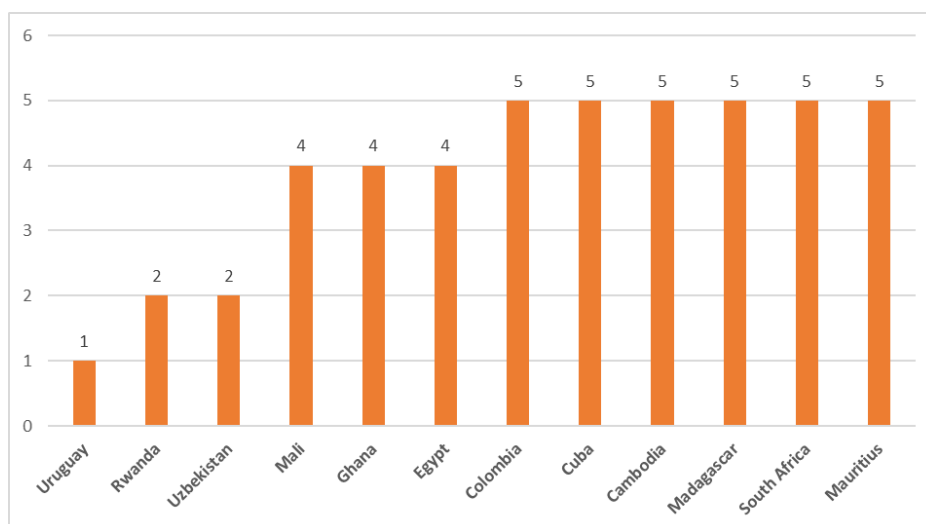
As for the recommendations, common weaknesses include the need for prioritization and/or clarity on who owns the recommended actions. There is also a need to specify actions that go beyond the project/programme boundaries to inform future CCA programming. This latter point needs to be taken with caution as CCA is highly context specific.

Example of a Satisfactory Report in Dimension 6 Conclusions, Lessons, Recommendations

The conclusions derived were complementary to the findings and the design of the project and gave further perspective to the ratings received for the indicators. The project was relevant, effective with appropriate level of efficiency. The design of the project was functional albeit with limitations such as lack of climate projections, promotion of alternate livelihoods, linking of coastal and terrestrial ecosystems, etc, which were factored into the analysis and results. The lessons were well founded with effective project actions and areas of opportunity which could have been merged with the recommendations. The lessons were also a difficult read because of how it was presented. However, the recommendations sections were very specific towards the project management unit, the executing entity, the implementing entity, and finally also AF. For the AF it suggested to use lessons and recommendations towards future projects while for the PMU, EE, and IE, there were actionable items such as conducting trainings and workshops, preparing integrated coastal management plan, etc. There was clear separation between these three sections.

Three reports had major (n=2) or severe (n=1) shortcomings. For these unsatisfactory reports, the main issues include the unclear and muddled conclusions, recommendations, and lessons and very limited lessons and recommendations generated in no small part due to weak analysis.

Figure 17 Ratings for Dimension 6 Conclusions, lessons, recommendations



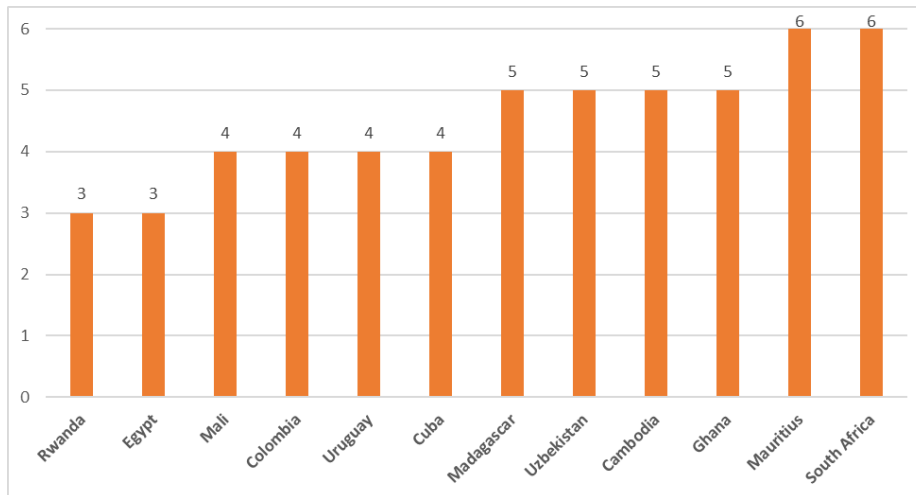
Dimension 7: Inclusion, ethics, independence

Dimension 7 assesses the extent to which the evaluations were undertaken in an inclusive, ethical, and independent manner. This means that the evaluation reports represented the views of a diverse range of stakeholders. In assessing against this quality dimension, the review also took account of the evaluation processes and outputs, and the evaluators' credentials. The AF works in adaptation where meaningful inclusion of different groups such as women, children, indigenous people, and others matter. Hence, the AF Evaluation Policy and relevant guidance notes along with the AF Gender Policy and Action Plan 2021 as expected are strong on this aspect. The review also considered the ethical standards that the evaluators adhered to as well as the cultural and technical competencies of the evaluators.

Ten evaluation reports were rated Moderately Satisfactory (n=4), Satisfactory (n=4), and Highly Satisfactory (n=2) thereby majority of the reports scored well in this dimension. Some of the MIE-commissioned evaluations included the sworn statement for ethical evaluation in the annex; others included a sub-section for this purpose.

Only two reports were Moderately Unsatisfactory, and the main weakness of these reports is that they did not elaborate on how the evaluation was independent, ethical, and inclusive. One of the two reports did not represent diverse voices in in the report even when the data collection clearly included a diverse range of stakeholders.

Figure 18 Ratings for Dimension 7 Inclusion, ethics, independence

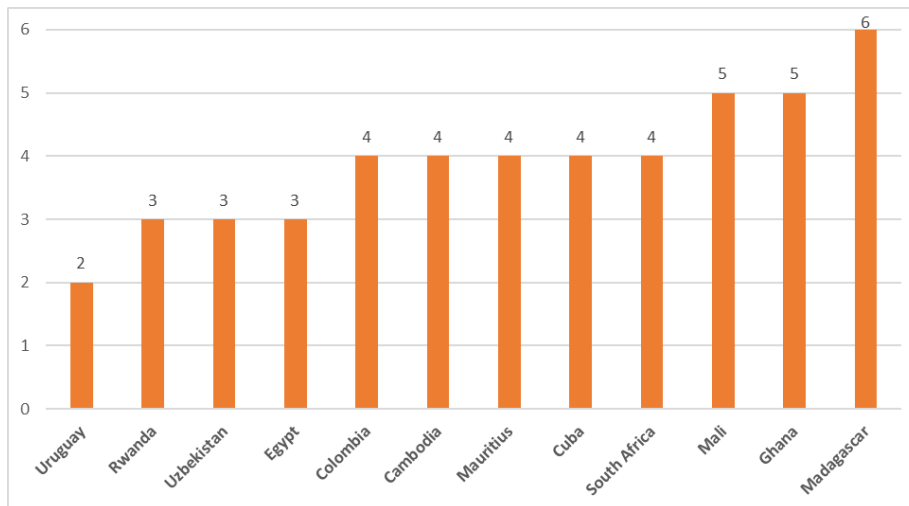


Dimension 8: Management and governance

Dimension 8 investigated the clarity of an evaluation’s management and governance, and in particular how these structures promoted the independence and credibility of the evaluation. This means that accountabilities, responsibilities, and lines of communication within an evaluation team, and between the evaluation team and IE were clear. It also entails appropriate provisions and timing for the quality assurance of the report.

Similar to the earlier trends, majority (n=8) of the evaluation reports were in the satisfactory region (Moderately Satisfactory = 5; Satisfactory = 2; and Highly Satisfactory = 1) and only had moderate to no shortcomings. There were four unsatisfactory (Moderately Unsatisfactory = 3; and Unsatisfactory = 1) reports.

Figure 19 Ratings for Dimension 8 Management and governance



Dimension 9: Utility

Dimension 9 assessed whether an evaluation report was written to meet the information and decision-making needs of the intended users and other stakeholders. In so doing, the review considered if an evaluation identified the potential users, how the report will be disseminated and communicated to them, the timeliness of the evaluation to inform decision-making, and the use of an evaluation management response to support evaluation follow-up.

Majority of the reports were either Moderately Unsatisfactory (n=5) or Moderately Satisfactory (n=4), leaving only three reports that are Unsatisfactory (n=1) and Satisfactory and Highly Satisfactory (n=2). The worst performing reports in this dimension did not identify the main users (e.g., primary, and secondary audience) of the evaluation and what follow up was to be done to ensure uptake of lessons and recommendations. In a couple of cases, the reports were generally of poor quality that they raise question on the reports' credibility, robustness, and hence usefulness.

Of the 11 reports whose report submission dates were available, two were not submitted within the required nine months of the project/programme closing dates stated in the reports. One of the two was over by about five months, and the other by a month. Interestingly, three reports were submitted before the closing date, which means they had been commissioned early in the final year of the project/programme. One report was submitted nearly six months before the recorded closing.

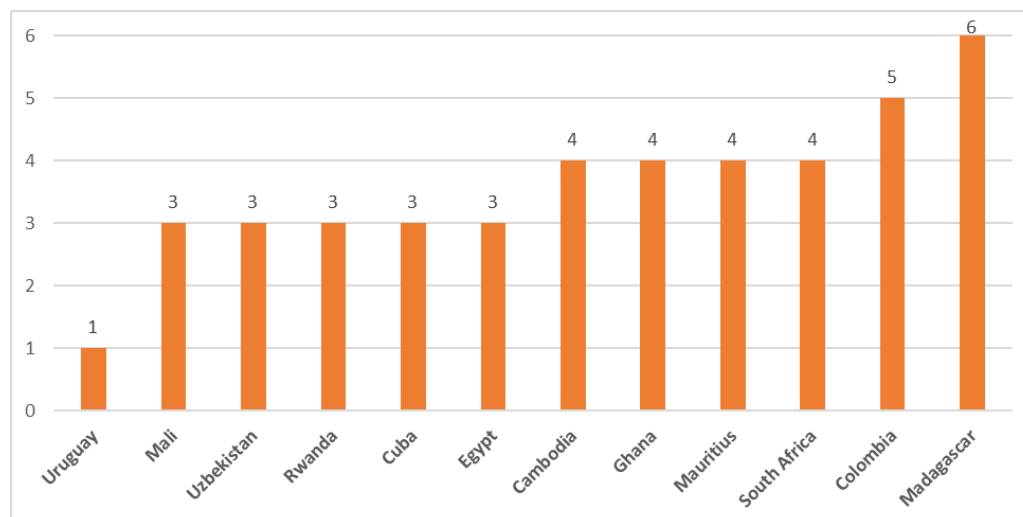
Example of a Highly Satisfactory Report in Dimension 9 Utility

The audience and users of the report were identified. A two-page brief was written and included in the report especially to improve utilization. The most important point the report noted is the need for translations to local language. The executive summary has French but the rest of the report, it was implied, might be translated as well. There was no management response to the report even though it was submitted at nine months after completion date.

Only two evaluation reports included management response. This, however, does not mean that there was no management response to these reports and such response did not feed into future programming. It could mean that the management response has not been included as part of the

report, or they have been submitted separately. In the case of some MIEs, management responses could be held in separate repositories. It is beyond the remit of this synthesis to search for these files and review them, which is essential to adequately respond to the fourth objective.

Figure 20 Ratings for Dimension 9 Utility



Section conclusion

This section presented the quality of the 12 final evaluation reports to respond to the first two synthesis questions. It did so by detailing the review findings in each quality dimension specified in the review methodology. In the process, the shortcomings of the reports against the quality dimensions were highlighted and examples of assessments provided. This conclusion brings the individual dimensions together and presents the overall rating based on the overall rating guide (Table 5) designed as part of the review methodology.

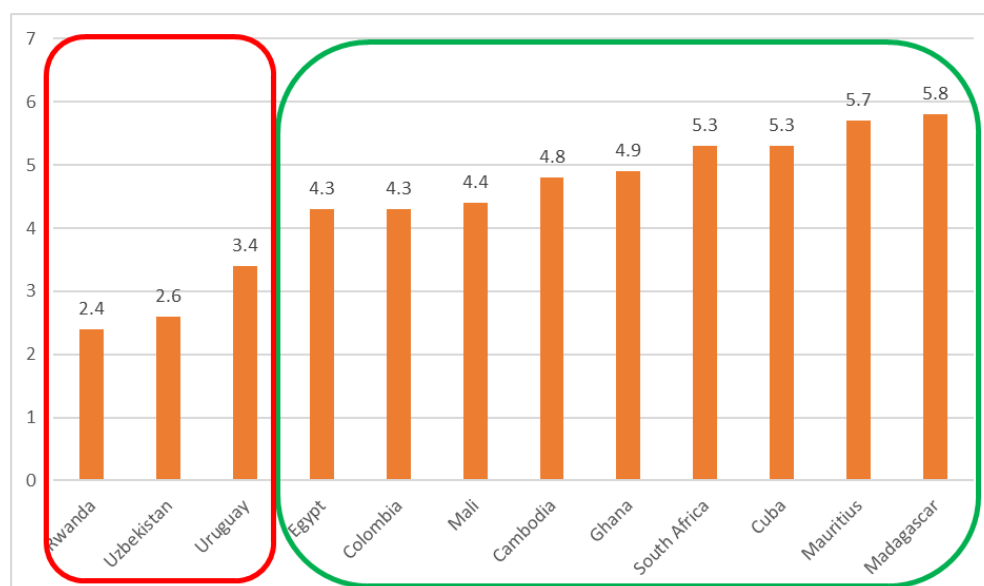
Table 7 Overall rating guide

Overall rating guide	
Highly Satisfactory: 5.5 - 6	The evaluation report fully meets or exceeds expected quality standards and there were no or few shortcomings. There is a high degree of confidence that the report is presenting robust evaluative findings, lessons, and recommendations.
Satisfactory: 4.8 - 5.49	The evaluation report meets expected quality standards with only minor shortcomings. There is confidence that the report can present robust evaluative findings, lessons, and recommendations when some minor improvements have been addressed.
Moderately Satisfactory: 4.3 - 4.79	The evaluation report meets only some of the expected quality standards with moderate shortcomings. The report can present robust evaluative findings, lessons, and recommendations when some moderate improvements are undertaken.
Moderately Unsatisfactory: 3.8 - 4.29	The evaluation report partially meets the expected quality standards but with noticeable shortcomings. The report can present robust evaluative findings, lessons, and recommendations when a good amount of improvements and corrections are undertaken.
Unsatisfactory: 2.3 - 3.79	The evaluation report only hardly meets expected quality standards with some shortcomings. The evaluation report requires substantive improvement before it can present robust evaluative findings, lessons, and recommendations.
Highly Unsatisfactory: 0 - 2.29	The evaluation report has major shortcomings in meeting expected quality standards. A major re-work is needed before it can present robust evaluative findings, lessons, and recommendations.

On balance, the quality of the portfolio of final evaluations included in this synthesis is satisfactory. The reports are largely compliant with the new AF Evaluation Policy and related guidance notes. Nine out of 12 reports were Moderately Satisfactory to Highly Satisfactory and only three were Unsatisfactory. Two of these three reports consistently did not score well in almost every quality dimension as demonstrated in the earlier assessment. It is important to note that two of the three reports were commissioned by an NIE, which points to future needs for evaluation capacity building. Among the most common shortcomings from the perspective of the AF Evaluation Policy include incomplete evaluation matrix, unsubstantiated findings, lessons that were phrased as

recommendations, limited lessons and recommendations beyond project/programme boundaries, unelaborated ethical standards, and management response that was not included in the report.

Figure 21 Overall quality rating of the final evaluation reports



The 12 evaluation reports were all utilized in this synthesis. As explained in Section 6 Methodology, this synthesis decided against an exclusion/inclusion threshold based on the quality of the reports. While it is common practice to exclude low quality evidence in a research or evaluation synthesis, the three unsatisfactory reports were included as important lessons can still be distilled from them. However, unsubstantiated lessons were systematically excluded from this synthesis.

9. Quality of projects/programmes

This section responds to synthesis question 3.

3. How well did the projects/programmes perform overall? Were they effective and if they are, are the results likely to be sustained? Are the longer-term results likely to be achieved? Did they perform better/worse than the portfolio of projects/programmes assessed in AF-TERG's first synthesis?

It largely addresses the following ToR objective:

Objective 1: To report to the Board on the overall performance and effectiveness of completed projects and programmes, including in relation to improved adaptive capacity and resilience.

This section provides an overall picture of the performance and effectiveness of the projects/programmes in the portfolio. The report draws on the findings from the 16 evaluations within the scope of AF-TERG's first synthesis to enable a comparative analysis against the ratings in the 11

evaluations covered by this synthesis. As stated in Section 6, one report from each synthesis²⁰ were not included.

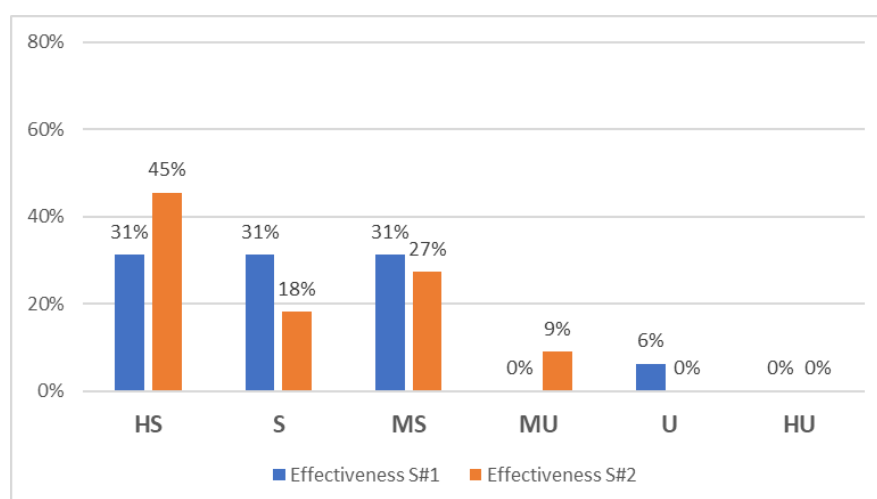
This section only included select dimensions – effectiveness, impact, and sustainability. The rest of the comparative analysis results have been interspersed in the discussion under Section 10.

Effectiveness

The trend in the effectiveness of the projects/programmes generally mirrors the overall ratings. A higher proportion of projects/programmes covered in the second synthesis is highly effective while more projects/programmes in the first synthesis are of satisfactory effectiveness. None of the projects/programmes was highly unsatisfactory, and the lowest effectiveness rating in this synthesis is moderately unsatisfactory. The diverse range of the intended outcomes of the projects/programmes in this synthesis is demonstrated in Annex 4. As expected, all the projects/programmes have an outcome that is related to increased resilience.

As in the overall rating, if the absence of the lowest ratings of unsatisfactory and highly unsatisfactory is to signify progress, then **the effectiveness of projects/progress in this synthesis paints a generally positive picture.**

Figure 22 Project/programme effectiveness ratings from the two syntheses

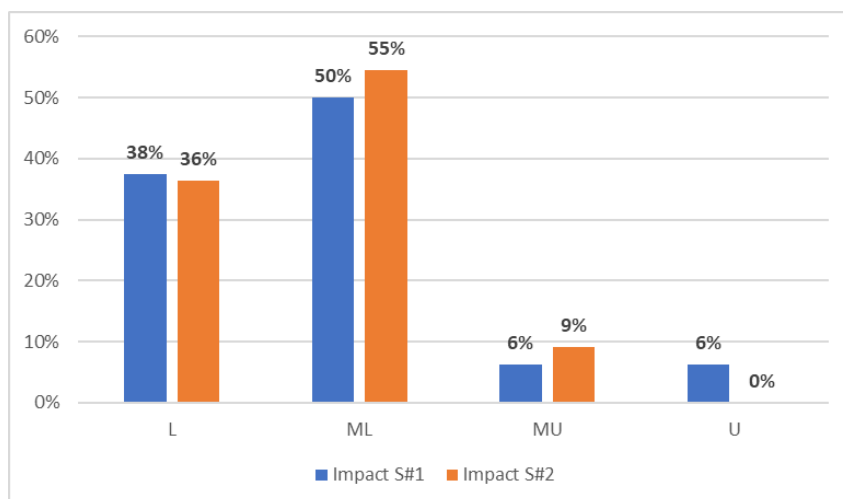


Likelihood of impact

More projects/programmes in the second synthesis are Moderately Likely to Likely achieve their intended impact. Akin to the overall and effectiveness ratings, there is a bit more spread in the performance of the projects/programmes in the first synthesis where a combined 12 per cent were rated Moderately Unlikely and Unlikely. There is an apparent absence of projects/programmes in the second synthesis at the bottom of the distribution and this is generally positive.

²⁰ The project in Cook Islands did not have an evaluation report, and the Rwanda evaluation did not include any rating rubrics.

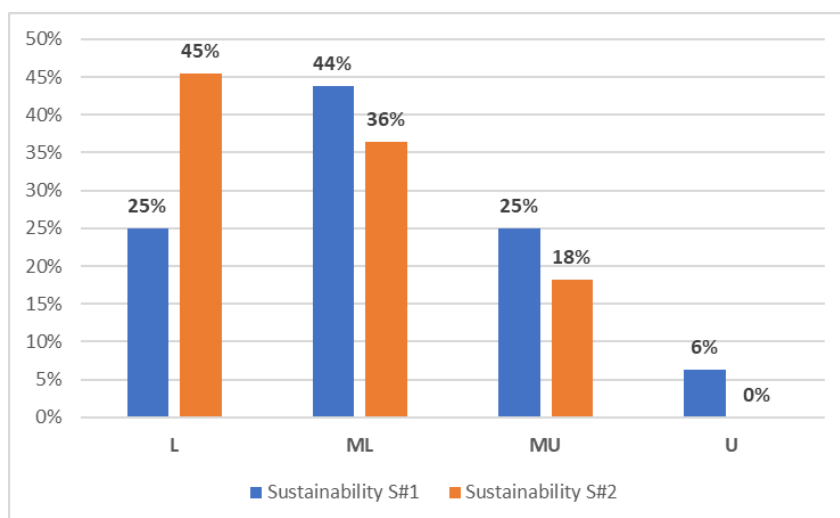
Figure 23 Project/programme impact ratings from the two syntheses



Estimated likelihood of sustainability

Over 80 per cent of the projects/programmes were estimated to moderately likely to likely sustain the benefits of the projects/programmes after their completion. This again is a positive finding especially since some of the operating contexts were difficult such as in Mali and northern Ghana. Over 30 per cent of the projects/programmes in the first synthesis were moderately unlikely to unlikely sustain the results, compared to 18 percent in the second synthesis.

Figure 24 Estimated project/programme likelihood sustainability ratings from the two syntheses

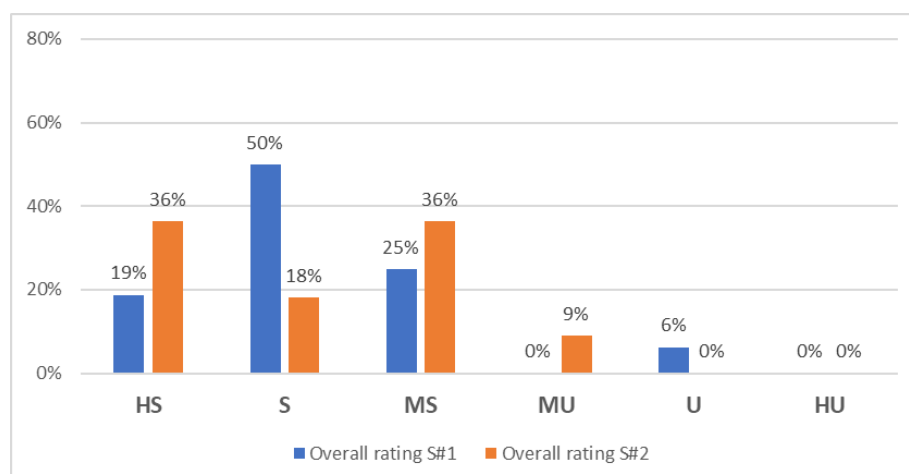


Section conclusion

It is unclear if the projects/programmes have improved. The proportion of projects/programmes that were rated overall Moderately Satisfactory to Highly Satisfactory were roughly the same for the two syntheses – around 94 per cent (n=15) for AF-TERG’s first synthesis compared to around 90 per cent (n=10) for this synthesis.

A higher proportion of projects/programmes (36 per cent) covered in this synthesis are Highly Satisfactory, compared to 19 per cent of projects/programmes covered in the first synthesis. Half of the projects/programmes in the first synthesis were Satisfactory.

Figure 25 Overall project/programme ratings from the two syntheses



While it is unclear if project/programme quality has improved, on balance, it can be concluded that the quality has not deteriorated. The lowest overall rating for projects/programmes in this synthesis is Moderately Unsatisfactory (9 per cent), compared to the first synthesis' Unsatisfactory (6 per cent).

10. Lessons from the synthesis

This section responds to synthesis questions 4 and 5.

It largely addresses **Objective 3**:

To assess the quality of the lessons in the final evaluation reports through the **identification of recurring findings, lessons learned and evidence on specific thematic areas of interest such as gender/equity, improved adaptive capacity and resilience, or sustainability.**

4. What are the common challenges (if any) affecting the performance across the portfolio of projects/programmes?
5. What are the common lessons that can be learned from the portfolio of projects/programmes?

When reading this section, readers will need to bear in mind that while the reports have identified rich findings and lessons many of them tend to be highly specific to the project/programme boundaries and local contexts. This is to be expected and does not automatically equate to poor lessons learning exercise by the evaluations. CCA projects/programmes respond to vulnerabilities that are specific to a given site and that are subject to specific climactic conditions and changes. Hence, findings and lessons that could potentially inform effectiveness and impact will most likely be context specific. Nevertheless, there were still some common themes that could be generated across the reports and this synthesis focused on the recurring challenges/issues and lessons learned.

Programmatic challenges and issues

Before discussing the common lessons learned from the 12 final evaluation reports, this synthesis will first present the most common (i.e., 50 per cent and above of the 12 reports) challenges that were

noted to have affected the performance and effectiveness of the projects/programmes. Understanding these challenges will put into context the lessons learned in the succeeding section.

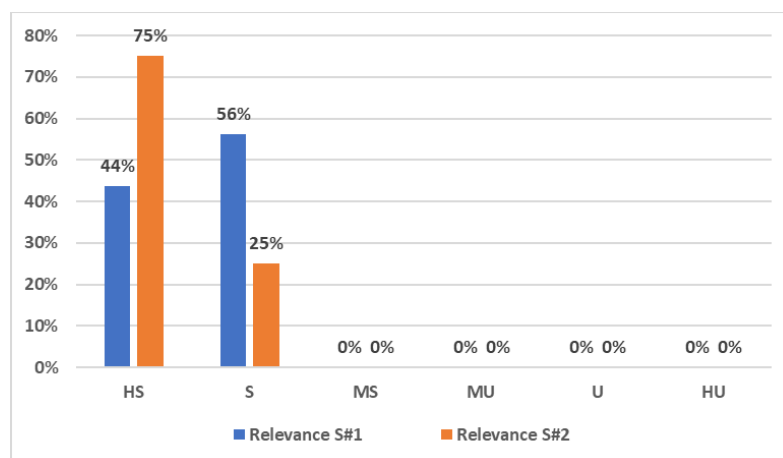
The top three most common challenges appearing in 11 of the 12 reports are as follows:

- **Flaws in the project design** including overambitious geographic scope and targets, lack of stakeholder analysis, needs analysis, and/or capacity gap analysis, and activities and outputs that do not have a clear link with the intended outcomes. One common challenge of projects/programmes with an alternative livelihood component is the absence of a value chain approach or commercial training for the beneficiaries, which affected both the viability and sustainability of results. Some issues centre around the alignment of support and community needs. There were also some issues with the mismatch between employment needs and the community needs.

These flaws are closely related to another recurring issue, the **lack of understanding of the local context** including of ecosystem, environment, and communities (n=6). Such a lack of understanding reportedly led to inappropriate outputs such as misplaced flood shelter, wrong plant species for replanting that ended up having very low survival rates, and others.

Despite these seemingly serious flaws, all the projects/programmes were generally satisfactory under the relevance criterion. In most of these cases, the overall criterion rating was due to averaging across a few dimensions within the relevant criterion instead of bad evaluation.

Figure 26 Relevance ratings from the first and second synthesis

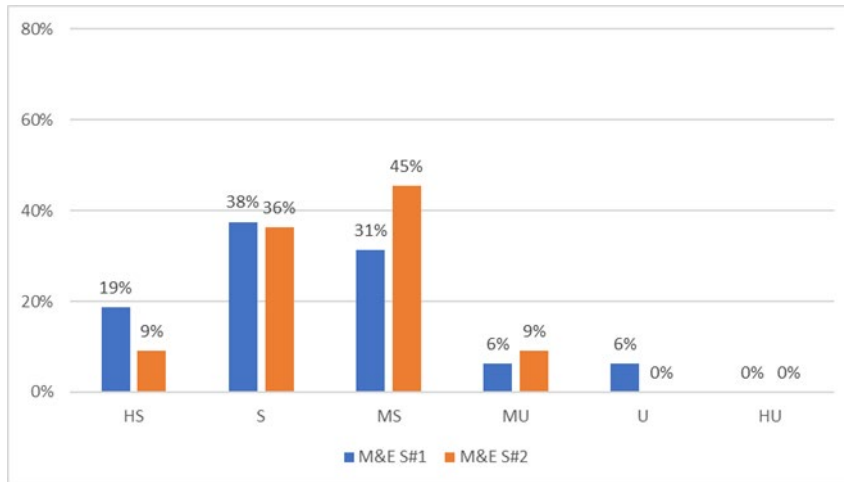


- **Operational and logistical issues** such as poor quality or high turnover of Implementing Entity staff, cumbersome procurement process, and the timing of release of funds. One report for instance cited a head of procurement who would require three quotations even for small purchases, which caused severe bottlenecks in the implementation. The slow release of funds came at different levels including from the Implementing Entity and the AF.
- **Low capacity** of implementing partners to successfully implement and manage projects/programmes. Weak capacities manifested in the lack of adequate qualified personnel, onerous or bad decision-making that led to delays, and poor reporting.

The fourth most common issue cited in 10 evaluation reports is **weak M&E system**, which hampered the ability of projects/programmes to learn, adapt, and improve its operations. Among the factors that contributed to this issue include poor overall design, lack of mechanism to measure improvements, delayed institutionalization of M&E, inadequate uptake and utilization of the framework and tools, and the inability to hire a dedicated M&E officer and undertake regular

monitoring activities due to minimal budget provisions for M&E. There was also reportedly an absence of M&E systems that can monitor and evaluate longer-term adaptation results. In some projects/programmes, challenges related to M&E were overcome by commissioning additional data collection to report at mid-term and this became the basis of the reporting thereafter. While this step is an established corrective action, it also needs to be noted that there are issues and challenges that come with retrospective data collection.

Figure 27 M&E ratings from the first and second synthesis

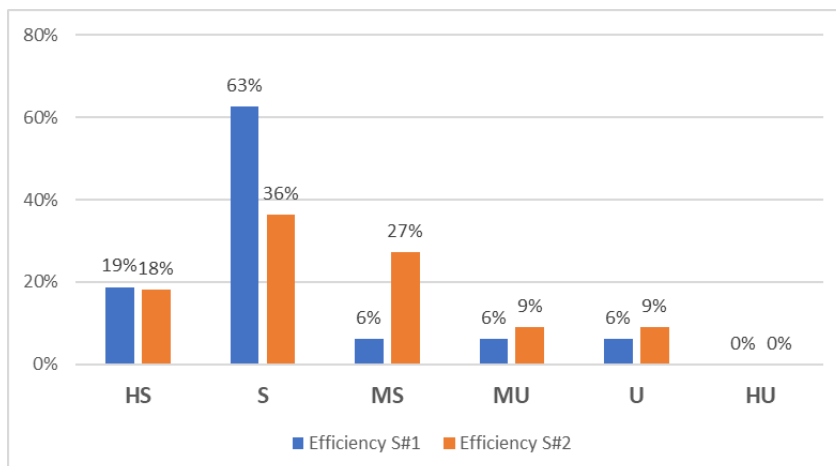


Other design-related issues cited include the **lack of** the following:

- **Beneficiary participation** leading to non-acceptance of new technologies, products, or livelihoods being introduced (n=8).
- **Gender and youth perspective** in planning and consequently in implementation (n=7). This meant that some activities such as training and outputs such as livelihoods have not adequately factored in the differing needs of various groups of beneficiaries. This also affected the ability of some projects/programmes to fully mobilize communities to achieve intended results since in some contexts the activities were met with resistance due to patriarchy.

It is worth noting that these most cited challenges all relate to project/programme relevance (design flaws) and efficiency (staffing, management, capacities, M&E).

Figure 28 Efficiency ratings from the first and second synthesis



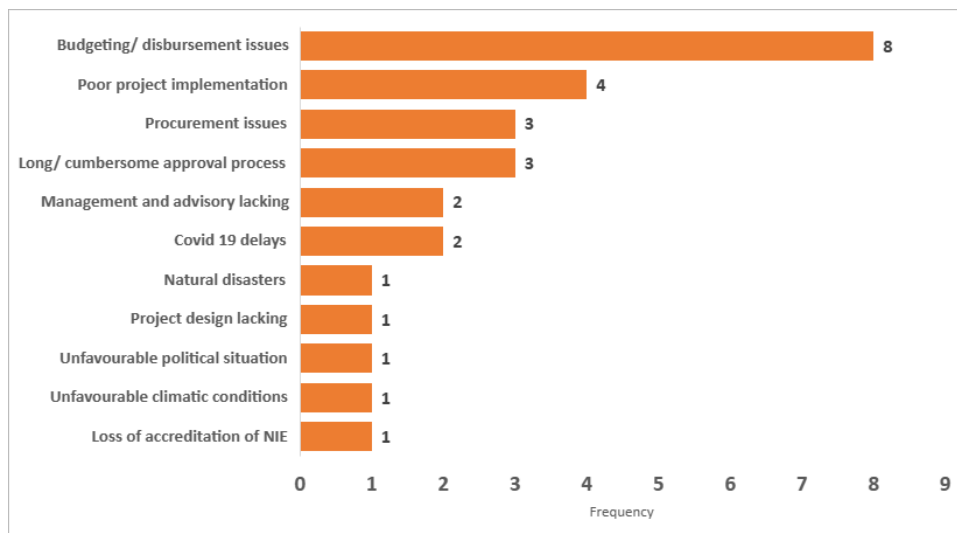
Despite these issues, the projects/programmes managed to achieve majority of their intended results leading to a generally high effectiveness ratings across the 12 projects/programmes.

In some cases, IEs and partners managed adaptively to overcome challenges. For example, in one case of delays in the release of funds, an IE pre-financed the mobilization to avoid further delays. There was also a case where procurement guidelines were changed to adapt to the capacities of the partners. An approach to overcoming design flaws adopted by a project/programme was to request for an extension to enable the project/programme to course correct and eventually achieve the intended results.

However, it is worth noting again that evaluation criteria ratings are based on the averaging of multiple dimensions and in the case of effectiveness, multiple results are considered. Evaluators make a judgement call as to how the overachievements fared despite the underachievement in some dimensions. Satisfactory rating can still be achieved even with underperformance in some respects.

Beyond the challenges within the control of the projects/programmes, there were also **external factors** that negatively affected performance and implementation period such as (i) disruptions due to Covid-19 (n=6) and (ii) natural disasters and security issues (n=8). The pandemic was noted to have caused some delays and/or procurement issues. Both natural disasters and security issues such as in Egypt, Ghana, Mali, and Uzbekistan to name a few also slowed down the implementation process. The most common reason for delays and thus extensions in the project implementation are in Figure 29.

Figure 29 Reasons for delays in project/programme implementation



Lessons from the portfolio

Four recurring lessons have been identified across the 12 evaluation reports. These lessons are related to the common challenges/issues identified earlier. Note that they also appeared as common lessons in the first synthesis and therefore it raises questions on how these earlier lessons were communicated to the IEs and how the IEs were incentivised to take up these lessons.

Lesson 1: The importance of understanding local context and environment (n=12). Such understanding and the overall knowledge of the ecosystem, political context, stakeholder and partner capacities, and community needs were noted to be highly pertinent to appropriate project design with fit-for-purpose components and implementation structure. They also facilitate adaptive management and inform decisions to enforce programmatic changes during implementation. Operating in difficult contexts such as those where security issues, political instability, natural disasters, and complex community dynamics prevail puts local knowledge front and centre of CCA programming.

Lesson 2: Inclusion is a key determinant of project/programme effectiveness and sustainability (n=11). The multi-dimensional needs and role of women, youth, indigenous people, and other groups must be built into the design, reviewed regularly, and mainstreamed into the implementation processes. A lack of focus on inclusion issues undermines the ability of projects/programmes to mark genuine positive impact on the adaptive capacity and resilience of individuals and communities.

Lesson 3: Strong partnerships and capacities are a pillar of successful implementation (n=10). Without the buy-in and participation of partners on the ground and relevant institutions and communities, project/programme efficiency suffer. While this may not necessarily and completely derail the achievement of intended results in every context as demonstrated by a few projects/programmes in this synthesis, the absence or limited partnership and capacity to deliver slows down the process of getting the much-needed adaptation funding to beneficiaries. And while capacity-building activities are a requirement in these contexts, genuine knowledge and information sharing at all levels should be the aim. This means that capacities and linkages will have to be built upward and downward in an often vertically structured government and funding bodies. As noted in one report, in CCA programming, it is important to do away with the assumption that grassroots organizations and local governments are the only ones that need to learn how to work with national and international organizations, when the latter often lack the capacity to engage locally.

Lesson 4: Robust M&E systems facilitate reporting, learning, and adaptive management (n=9). However, gaps remain in the specification of targets and indicators, utilization, and generally institutionalization/mainstreaming of a plan that broadly serves adaptation programming needs. Apart from adequate budget provisions within the project/programme time frame, M&E systems that have a longer time horizon are needed to capture adaptation results that only manifest after completion. This is true for projects/programmes that include reforestation, mangroves planting, livelihoods, and resettlement projects. While many of the evaluations in this synthesis attempted to evaluate progress in achieving outcomes and in one case with a strong theory-based orientation, the lack of data and timing of the final evaluation were noted to negatively affect these efforts.

11. Recommendations

This synthesis puts forward two sets of recommendations. One set relates to CCA programming, and the other set complies with the ToR's requirement to specify recommendations to improve the quality review methodology. The second set, along with the other recommendations that are internal to TERG, were redacted in this public version.

Programmatic recommendations for the AF Secretariat

These programmatic recommendations for the AF Secretariat focus on areas where limited improvements have been made based on the findings of the two syntheses. Given their persistence, response and actions to address them are urgent.

Recommendation 1: Strengthen due diligence of requirements that can foster better understanding of local contexts and environment during project development phase, including stakeholder analysis, beneficiary needs analysis, capacity gap analysis, and gender and other inclusion studies. They must not be treated as stand-alone requirements. The Secretariat check list during origination can include the clear link between the findings coming out of these analyses and the overall theory of change both in narrative and diagrammatic formats, implementation structures, as well as risk management mechanisms. The theory of change's components (i.e., risks and assumptions, outputs, outcomes, and objectives) themselves should take account of the findings of

the previously mentioned analyses. This will stand the IEs in a good stead to respond to the required changes to improve operations down the line.

Recommendation 2: Check for the adequacy of the M&E budget vis-à-vis M&E plans at entry and require updates on how M&E is working (or not) in progress reports. Weak M&E was a recurring issue among the projects/programmes in this synthesis. But there are many facets to an M&E system and there won't be a one-size-fits-all solution to every issue that contributes to quality shortfalls. There are also external factors that could affect the efficiency and effectiveness of a system during implementation. Hence, a solution is to check that there is a sound M&E plan as well as adequate budget that can support the components of the plan during the design stage. Thereafter, a short section in the progress reporting template could be dedicated on assessing how the M&E is operating. One project/programme covered in this synthesis did not have any M&E in place until the mid-term review process and another one did not have any dedicated M&E officer and management information system at all. These situations could have been avoided had there been checks during implementation.

Recommendation 3: Conduct a benchmarking exercise on the scale of M&E budgets in approved proposals in collaboration with AF-TERG. Given the lack of information in the FE reports, it will be important to understand the scale of the M&E budget of AF-funded projects/programmes in various contexts and sectors. Follow up interviews with the IEs can then be undertaken to complement the benchmarking exercise and test the adequacy of various budgetary range. Findings from this exercise can inform future guidance either from AF Board Secretariat or AF-TERG on budgeting for both project/programme monitoring and evaluations that is more specific to the AF operating context.

Annexes



TERMS OF REFERENCE

Short-Term Consultant (STC) to Support Synthesis of Adaptation Fund Final Evaluations (Part 2)

Introduction

The purpose of these terms of reference (ToR) is to elaborate the requirements for the support of the proposed Synthesis of Adaptation Fund Final Evaluations (FEs) covering the period from July 2020 to August 2023. More specifically, it aims to provide key information to stakeholders about the synthesis exercise, and to guide the consultant, including on the specific requirements during the synthesis process. The ToR are structured as follows: section 1 presents the background to the synthesis of FEs; section 2 presents the rationale, objectives, stakeholders, and main users of the synthesis; section 3 presents the purpose and objectives of the second Final Evaluation Synthesis; section 4 identifies the approach and methodology; and section 5 outlines the timelines and key milestones. The required experience, qualifications and competencies of the consultant are included in section 6.

1. Background

The Adaptation Fund was established by UNFCCC to finance concrete adaptation projects and programmes in developing countries. Since 2010, the Adaptation Fund (AF) has committed over US\$ 1 billion for climate change adaptation (CCA) and resilience projects and programmes, including 150 concrete, localized projects in the most vulnerable communities of developing countries around the world with over 38 million total beneficiaries. The Fund is financed largely by government and private contributors, and from a two per cent share of proceeds of Certified Emission Reductions (CERs) issued under the UNFCCC Kyoto Protocol's Clean Development Mechanism projects.

The Technical Evaluation Reference Group of the Adaptation Fund (AF-TERG) was established in 2018, as an independent evaluation advisory group accountable to the Adaptation Fund Board (the 'AF Board' or the 'Board'), to ensure the independent implementation of the Fund's evaluation framework, which was replaced by the Evaluation Policy (2022)²¹ in October 2023. As per its multi-year work programme, the AF-TERG committed to undertake independent evaluations according to priority evaluation needs, including annual evaluation syntheses from final evaluations (FE) of Adaptation Fund's projects.

Paragraph 66 of the AF Operational Policies and Guidelines for Parties to Access Resources from the Adaptation Fund²² states that "[a]ll regular size concrete projects and programmes that complete

²¹ <https://www.adaptation-fund.org/wp-content/uploads/2022/07/Evaluation-Policy-of-the-Adaptation-Fund.pdf>

²² https://www.adaptation-fund.org/wp-content/uploads/2017/08/Amended-OPG_Oct-2022_2.pdf

implementation will be subject to terminal evaluation by an independent evaluator selected by the Implementing Entity. All small size concrete projects and programmes, as well as readiness grant projects, shall be subject to terminal evaluation if deemed appropriate by the Board and shall follow an evaluation process as decided by the Board using templates approved by the Board. Terminal evaluation reports will be submitted to the Board as stipulated in the project agreement.” The Evaluation Policy further reaffirmed this by mandating that “[a]ll Fund-supported projects and programmes that complete implementation should conduct a final evaluation to assess project/programme performance and impact to support learning and accountability, and inform future CCA interventions... All IEs are required to commission an independent final evaluation of their projects, submitted to the secretariat and the DA within nine months of project completion.” The Evaluation Policy goes even further by requiring IEs to submit management response that... “[d]escribes what, why, and how final evaluation learning will be incorporated into current or future AF Fund work.”

In this way, terminal evaluation has been made a core element of the learning and accountability of the AF. It is also important to note that terminal evaluations are crucial particularly for adaptation projects where measurement of results, including impacts, is a more complex process and learning is of particular importance to capture the approaches that work and to provide input into the future programming of the AF, and adaptation finance as a whole.

Evaluation syntheses therefore intend to inform the work of the AF and partners by synthesizing evaluative evidence, lessons learned, and conclusions from evaluations of projects funded by the Fund and co-financed by other relevant institutions. They are also intended to reflect on the quality of the final evaluations, as well as identify capacity gaps, and make recommendations to the Board for improvement so that there is a consistent improvement in the standard and usefulness of these evaluations to the Fund and its stakeholders.

It is important to note that this is a transition phase, as the evaluation subjects to this synthesis were not prepared under the Evaluation Policy, while the synthesis will be required to apply the requirements of the policy. However, such an approach will allow to identify gaps in the evaluation quality and capacity against the Evaluation Policy and provide, among others, feedback on the need for further capacity building and development of guidance and support materials to IEs.

The Evaluation Policy seeks to promote (i) accountability for the achievement of the Fund objectives; and (ii) learning, feedback, and knowledge-sharing on results and lessons learned among different groups participating in the Fund. It also establishes requirements for how the Fund activities should be evaluated in line with international principles, norms, and standards and sets up the Fund’s evaluation principles and criteria. It is intended to add value and contribute towards the achievement of the Fund’s goal, and to the realization of the planned social and environmental impact.

2. Context of Final Evaluations Synthesis

The development of evaluation syntheses of AF final evaluations is fully aligned with the requirements of the Adaptation Fund Medium Term Strategy 2023-2027 (MTS) and intends to inform the future work of the Fund and its partners by feeding it evaluative evidence, lessons learned and conclusions from evaluations of completed adaptation projects.

In its indicative work programme for fiscal years 2021 to 2023 (FY21 – FY23)²³, AF-TERG seeks to generate evaluative insights and knowledge specifically by articulating and utilizing evaluation results and learning within the Fund. The initial evaluation synthesis conducted in 2020 was the first of a

²³ Fiscal Years 2023 – 2024 Update To The Work Programme Of The Adaptation Fund Technical Evaluation Reference Group (AF-TERG) (https://www.adaptation-fund.org/wp-content/uploads/2022/03/AFB.EFC_.29.7-AF-TERG-Work-Programme-FY23-24-Update_final.pdf)

series of evaluation syntheses to be produced by AF-TERG, as part of its initial work progress in FY20 and in accordance with the technical group's objectives. The initial synthesis presented the first quality review of AF final evaluations reports. That synthesis covered all 17 final evaluations of projects/programmes received from 2015 to June 2020 and was, therefore, based on the first cohort of completed projects in the Fund portfolio. Of the 15 projects/programmes by Multilateral Implementing Entities (MIEs), 12 were implemented by the United Nations Development Programme (UNDP).

The present second synthesis aims to expand the learning from final evaluations of adaptation projects/programmes and strengthen AF TERG's efforts to provide synthesis products of interest to different target audiences within AF. It is timed to inform the formulation and implementation of the AF-TERG's new multi-year work programme, further the work of the AF and its Implementing Entities and, above all, to provide an important frame of reference for the AF Evaluation Policy.

3. Purpose and objectives of the second final evaluation synthesis

Purpose. The main purpose of the second synthesis of final evaluations is to establish an overall methodology for reviewing of final evaluations leveraging on the quality rating criteria established by the AF in 2015²⁴. This process involves assessing the quality of the new set of final evaluations of AF projects/programmes. By doing so, we aim to provide the Adaptation Fund stakeholders with detailed insights into various aspects such as project performance, learning experiences, implementation strategies, and the extent to which the final evaluation findings have been utilized. This periodic exercise facilitates a structured reporting to the Board, presenting an aggregated overview of the quality inherent in the final evaluation reports, alongside an analysis of the performance encompassed within the AF portfolio at the project's completion stage.

Objectives. The main objectives of the synthesis are:

1. To report to the Board on the quality of final evaluations and on the overall performance and effectiveness of completed projects and programmes, including in relation to improved adaptive capacity and resilience.
2. To review the quality of final evaluations reports based on compliance with the criteria as laid out in the Evaluation Policy and Guidance of Final Evaluations.
3. To assess the quality of the lessons in the final evaluation reports through the identification of recurring findings, lessons learned and evidence on specific thematic areas of interest such as gender/equity, improved adaptive capacity and resilience, or sustainability.
4. Serve an accountability purpose through exploring the extent of management response to recommendations provided in the Final Evaluations and how the management response is fed into future programming.

4. Methodology

The second final evaluations synthesis will present a quality assessment of the portfolio of final evaluations and provide an overview of results from the evaluations of completed Adaptation Fund projects. The quality review will inform the lessons and implications for future final evaluations reports, with a specific attention to good evaluation practices. Additionally, this synthesis is intended to establish a systematic quality assurance methodology for final evaluations. Therefore, part of this work will involve the development and piloting of such methodology. The steps to this synthesis are summarized below:

²⁴Guidelines for project/programme final evaluations.

https://www.adaptation-fund.org/wp-content/uploads/2015/01/Guidelines%20for%20Proj_Prog%20Final%20Evaluations%20final%20compressed.pdf

Step 1: Development of a draft methodology for the review of AF final evaluations

The consultant is encouraged to draw upon existing approaches, including reference materials such as the GEF IEO “Methodological Paper #9: Approach to Terminal Evaluation (2009)”. While developing the methodology, it is imperative that the consultant ensures its full alignment with the Evaluation Policy and other relevant documents of the Fund.

The methodology may also cover some of the questions below:

Criteria	Evaluation Questions
Report quality	1. To what extent do final evaluation reports meet the evaluation criteria ²⁵ to assess quality? Is a management response provided? Have adequate resources been used?
Quality of evidence	2. What is the quality of the evidence and methodologies used by final evaluations? To what extent are the methodologies and evidence appropriate, and the ratings well substantiated? Did the FEs explore the quality of M&E at entry and during implementation?
Quality of Lessons identified	3. What is the quality of the lessons identified in the FE reports? Do final evaluations provide the evidence to substantiate these lessons?
Evaluation of risks to sustainability	4. What is the likelihood that the achieved and expected outcomes (including in relation to climate change adaptation and resilience) would be sustained in the portfolio of completed projects?
Contribution to Equity/Gender equality	5. What is the quality of supporting processes for promotion of equity and gender equality in the portfolio of completed projects? To what extent do they meet the guidance provided in the Guidance document for Implementing Entities on compliance with the 2016 Adaptation Fund Environmental and Social Policy ²⁶ and the AF Gender Policy and Gender Action Plan approved in 2021 ²⁷

The output of this step will be a final draft methodology. The final draft methodology shall be reviewed and approved by the AF-TERG before applying it in the next step.

Step 2: Application of the methodology to the second cohort of final evaluations

The evaluation synthesis will assess final evaluations submitted to the Adaptation Fund from June 2020 to August 2023, i.e. the dates at which the first and latest evaluations reports were received. The initial scoping exercise shows that there are currently 11 FEs that will form part of this synthesis. The total numbers of FEs will be confirmed during inception.

Since the study intends to assess the reporting processes and not the projects themselves, it will be entirely desk-based and will utilize mainly the information presented in the final evaluation reports. It is planned that the AF-TERG will use the final evaluations as key and essential inputs (based on those evaluations that the current study will identify as having an acceptable quality) to future evaluations. The synthesis will be based on the Evaluation Policy Guidelines and the methodology to be developed will be applied to the set of guidelines and criteria provided in the new policy.

The output of this step will be a draft synthesis report applying the draft methodology, as well as recommendations for refinement of the draft methodology.

²⁵ An essential part of each final evaluation is the assessment of the project/programme relative to the Fund’s nine evaluation criteria as outlined in the Evaluation Policy (2022): Relevance; Coherence; Effectiveness; Efficiency; Impact; Equity; Adaptive management; Scalability; Human and ecological sustainability and security.

²⁶https://www.adaptation-fund.org/wp-content/uploads/2016/07/ESP-Guidance_Revised-in-June-2016_Guidance-document-for-Implementing-Entities-on-compliance-with-the-Adaptation-Fund-Environmental-and-Social-Policy.pdf3

²⁷https://www.adaptation-fund.org/wp-content/uploads/2016/04/OPG-Annex-4_GP-and-GAP_approved-March2021pdf-1.pdf

Step 3: Final synthesis report and final methodology

In this step, the consultant will finalize the report from Step 2 and the methodology from Step 1, following inputs from members of the AF-TERG, and other relevant stakeholders, such as AFB Secretariat.

The output of this step will be a final synthesis report and a methodology for review of final evaluations, both cleared by the AF-TERG.

5. Organization and schedule

The consultant will closely work with one of the members of the AF-TERG who will be appointed as a focal point for this evaluation synthesis. The focal point will provide strategic guidance and quality assurance of all final deliverables of this assignment. The consultant will report to the AF-TERG Secretariat Coordinator for administrative matters.

The assignment was delivered within 30 expert days in the period October 16 – December 15, 2023. The project milestone and dates are indicated in the table below.

Project milestone	Completion Date
Draft evaluation review methodology	October 30, 2023
Draft synthesisist report of final evaluations and comments on improvement of the methodology	November 27, 2023
Final report and methodology	December 15, 2023

6. Experience and qualifications

Education: Strong and proven academic background in applied research and climate change sciences, with a PhD /master's level degree or equivalent in a relevant field (climate change adaptation; international development; social sciences or related area).

Experience:

- Minimum of 10 years of experience in leading and implementing evaluations or reviews in diverse climate change adaptation, resilience and development contexts.
- Experience in reviews or evaluations of projects of the Adaptation Fund projects or any other projects of other climate finance organizations.
- Expertise and track record of reviewing or evaluating complex projects
- Ability to communicate complex technical detail with clarity, and ability to identify and focus on the key messages among competing detail.
- Strong interpersonal skills, ability to work with different stakeholders and skills to promote good communication and a learning culture.

Evaluation criteria

Criteria	Points	Percentage
Qualification		20%
Strong and proven academic background in applied research and climate change sciences, with a PhD /Master's level degree or equivalent in a relevant field (climate change adaptation; international development; social sciences or related area).	20	
Experience		70%
Minimum of 10 years of experience in leading and implementing evaluations or reviews in diverse climate change adaptation, resilience and development contexts.	20	

Experience in reviews or evaluations of projects of the Adaptation Fund projects or any other projects of other climate fund organizations.	40	
Ability to communicate complex technical detail with clarity, and ability to identify and focus on the key messages among competing detail.	10	
Competences		10%
High level planning, organisational and time management skills, including flexibility, attention to detail and the ability to work under pressure to meet challenging deadlines;	5	
Excellent interpersonal skills	5	
Total		100%

Annex 2 Final evaluation quality review framework

Explanatory note to the draft evaluation review methodology

Eunica Aure

30 October 2022

I. Introduction

The Adaptation Fund Technical Evaluation Reference Group (AF-TERG) commissioned a short-term consultant to undertake a synthesis of the AF's final evaluations submitted between 2020 and 2023. The assignment started on 16 October 2022. It is comprised of three deliverables: (i) a draft evaluation review methodology, (ii) a draft evaluation synthesis report and improvements to the review methodology, and (iii) final report and methodology.

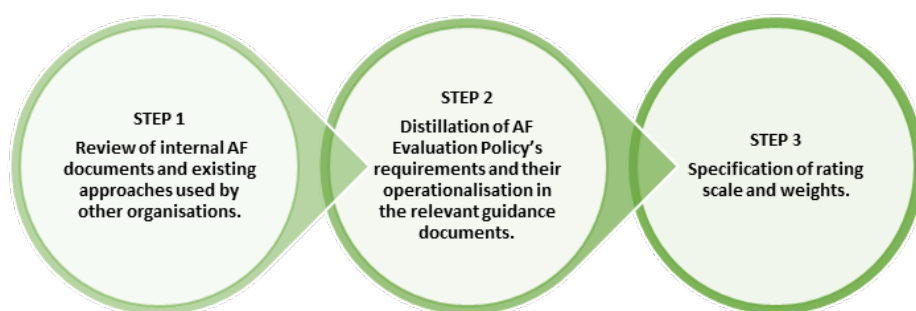
This explanatory note sets out the process through which the first deliverable, the draft evaluation review methodology, was formulated. The note must be read alongside the draft methodology in the accompanying Excel spreadsheet.

As agreed during the AF-TERG meeting on 16 October 2023, this review methodology forms part of the synthesis work to assess the quality of the final evaluations. This draft methodology is distinct from the framework to be used for the synthesis itself, which will focus on identifying and capturing emerging lessons to inform future decision-making.

II. Approach to the development of the draft evaluation review methodology

The terms of reference (ToR) states that the review methodology will 'draw upon existing approaches'. The methodology shall be fully aligned with the AF Evaluation Policy, its relevant Guidance notes, and other relevant documents of the Fund. To learn from how other organizations are undertaking this type of quality assurance or assessment of evaluation reports, the consultant also considered existing approaches and frameworks where they were publicly available. The process of developing the draft methodology followed three steps.

Figure 30 Three-step approach to the development of the draft methodology



III. The draft evaluation review methodology

The draft methodology is presented in what follows under the three-step approach.

Step 1: Review of internal AF documents and existing approaches used by other organizations.

The consultant reviewed the new AF Evaluation Policy along with the Guidance²⁸ documents, the AF's Evaluation Framework and Guidelines for Project/Programme Evaluations, and the 2021 Synthesis of Final Evaluations report.

Existing approaches reviewed include [GEF IEO's Guidelines for Conducting Terminal Evaluations of Full-Size Projects](#); the GCF Secretariat's Evaluation Quality Assurance and Review Framework for AE-led Evaluations in [Evaluation Operational Procedures and Guidelines](#); [WFP's Post-hoc Quality Assessment for Evaluations \(PHQA\)](#), [UNFPA's Evaluation Quality Assurance and Assessment \(EQAA\)](#), and UK FCDO's [Evaluation Quality Assurance and Learning Services \(EQuALS\)](#). WFP also has an [Evaluation Quality Assurance System \(EQAS\)](#), which provides quality standards and guidance on evaluations managed by the Office of Evaluation and/or by country offices, regional bureaus, and headquarters divisions.

It is important to note that most of these existing approaches are applied to individual evaluation reports upon submission as part of a quality assurance process, as opposed to a group of evaluations for quality assessment and synthesis purposes as in this assignment. There is also a difference on who performs the reviews. For example, the GCF Secretariat assess all mid-term and final evaluations through internal staff resources. The UK FCDO and WFP outsource the function to a firm or a consortium of firms.

Step 2: Distillation of AF Evaluation Policy's requirements and their operationalization in the relevant guidance documents.

The AF Evaluation Policy requires that all of its evaluation principles to apply at different stages of evaluation processes and inform the ToR, selection of evaluators, manner in which evaluations are conducted, **assessment of evaluation products**, management responses, and communication and utilization.²⁹ Any exception in the application of the principles must be justified to the AF-TERG.³⁰ As such, the AF evaluation principles provide a clear backbone to the proposed evaluation review methodology.

The AF Evaluation Policy has seven evaluation principles:

1. Relevance and utility
2. Credibility and robustness
3. Transparency
4. Impartiality and objectivity
5. Equitable and gender-sensitivity inclusivity
6. Complementarity
7. Complexity-sensitive and adaptive

These principles were translated into quality dimensions that form the organizing structure of the proposed methodology. Under each principle, the precise requirements of the AF Evaluation Policy and their operationalization in the guidance notes were generated. For example, under the principle of Relevance and Utility, the guidance note for Evaluation Principles specified the need to identify how the evaluation findings, lessons, and recommendations will be communicated to intended users.³¹ The guidance for Evaluation Reporting further requires report formats to be practical,

²⁸ Such as the Guidance supporting the operationalization of the (i) Evaluation Principles; (ii) Evaluation Criteria; (iii) Evaluation Budgeting; (iv) Evaluation Reporting; and (v) Final Evaluations.

²⁹ p.12, AF Evaluation Policy

³⁰ p.11, AF Evaluation Policy

³¹ p. 6 and p. 12, Guidance in Support of the Operationalization of the Evaluation Policy: Evaluations Principles

concise, user-friendly, readable, and accessible to the intended audience.³² For this reporting requirement, the consultant made the judgement on its relationship to the principle of utility – the more user-friendly an output is, the more likely it will be utilized by intended users. The final output of this mapping exercise is a long list of quality dimensions for each evaluation principle.

As there are different ways to group the dimensions, the consultant provided three options to the AF-TERG focal point and AF-TERG Secretariat on the logical grouping: by principle, by evaluation report section, and by evaluation elements. The options brief (not a formal deliverable) is in Annex 1. After this correspondence, it was decided that the final grouping for the draft methodology will inevitably be a mix of evaluation elements, evaluation report sections, and other high-level headings (e.g., structure and clarity, inclusion, ethics, and independence, and others). Nine quality dimensions were selected in the end to represent comprehensively the requirements of the AF Evaluation Policy.

1. Structure and clarity
2. Context, purpose, scope, and objectives
3. Evaluation framework and methodology
4. Methods and data
5. Analysis and findings
6. Conclusions, lessons, and recommendations
7. Inclusion, ethics, and independence
8. Management and governance
9. Utility

These dimensions were then broken down into more specific considerations (please refer to the Excel spreadsheet) both for transparency in the assessment process and to provide guidance to future reviewers.

Note that the draft methodology aimed to be comprehensive in covering the requirements of the AF Evaluation Policy. **There is scope to streamline the dimensions and specific considerations depending on the priorities of the AF-TERG.**

Step 3: Specification of rating scale and weights.

The AF has a 'mandatory rubrics rating scale' that final evaluations are required to apply to assess and communicate the performance of a project/programme against the Evaluation Policy's evaluation criteria.³³ The rubrics use either six or four rating levels as per the Guidance document, with no requirement on the use of one over the other.³⁴

The draft methodology proposes to similarly adopt an even numbered Likert scale with six rating levels.

³² p.13, Guidance in Support of the Operationalization of the Evaluation Policy: Evaluations Reporting

³³ p. 15, Guidance in Support of the Operationalization of the Evaluation Policy: Final Evaluations

³⁴ *ibid.*

Table 8 Proposed rating scale in the evaluation review methodology

Individual criterion rating guide	Rating	Description	Explanation
	6	Highly Satisfactory	The criteria were fully met or exceeded and there were no shortcomings.
	5	Satisfactory	The criteria were met with only minor shortcomings.
	4	Moderately Satisfactory	The criteria were partially met with moderate shortcomings.
	3	Moderately Unsatisfactory	The criteria were partially met with noticeable shortcomings.
	2	Unsatisfactory	The criteria were somewhat met with major shortcomings.
	1	Highly Unsatisfactory	There were severe shortcomings in meeting expected standards.

The decision between four or six levels is pragmatic than technical. If simplicity and parsimony are priorities, four rating levels may be more appropriate. However, if AF-TERG prefers to see more gradient in the distribution of quality, the six rating levels may be used as it was adopted in the preceding 2021 synthesis.

Each of the quality dimension ('criterion') will be rated 1 to 6 depending on the extent of the evaluation reports' shortcomings in meeting the criterion. The overall score can be calculated using the weights proposed in Table 2. The weights reflect the level of importance of each criterion in the assessment. **Similar to the rating scale, the weights can be redistributed depending on AF-TERG's ordering of the dimensions' importance.**

In the proposed weights (Table 2), the biggest importance is accorded to Methods and data, Analysis and findings, and Conclusions, lessons, and recommendations with 20 points each. The robustness of methods and data, and how they were used in the analysis and to inform the lessons and recommendations are highly pertinent to an evaluation's usefulness and credibility; they are the practical application of the Methodology and framework (10 points) and Inclusion, ethics, and independence (10 points). The other dimensions were assigned five points each as compared the other dimensions, their content is often readily available but is a matter of describing them properly in an evaluation report.

Table 9 Proposed weighting for each criterion in the evaluation review methodology

Criteria	Weighting
1. Structure and clarity	5
2. Context, purpose, scope, and objectives	5
3. Evaluation framework and methodology	10
4. Methods and data	20
5. Analysis and findings	20
6. Conclusions, lessons, and recommendations	20
7. Inclusions, ethics, and independence	10
8. Management and governance	5
9. Utility	5
TOTAL WEIGHTED SCORE	100

The full set of quality dimensions and specific recommendations, the rating scale, and the weights together make up the template in the Excel spreadsheet. This template was developed to facilitate the assessment in the most transparent manner.

Assessment criteria and sub-criteria

1. **Structure and Clarity** - The evaluation report is logically structured, accessible, and contains all relevant elements of an evaluation report.
 - The evaluation report structure is logical (e.g. logical use of sections and sub-sections).
 - The evaluation report is written in a coherent and accessible manner including to the local population (e.g. free of jargon, written in plain English, appropriate use of tables, graphs and diagrams, all abbreviations explained, total number of pages is not overly long).
 - A concise executive summary is included, and it provides an accurate summary of the main product (as opposed to a full copy and paste of sections).
 - Where annexes are provided, they are useful, relevant, and necessary to support evaluation findings and analysis.

2. **Context, Purpose, Scope, and Objectives** - There is a sufficiently detailed description of the background to the evaluation, including the context, purpose, scope, and objectives.
 - The evaluation report provides a clear but succinct description of the project to be evaluated, including sufficient understanding of the role of the project within a given system and its linkages with specific organizations, sectors, thematic areas, or geographic space to support broad complementary learning (rather than an exclusive intervention – or institution-centric perspective).
 - The evaluation report provides a relevant and sufficient description of whether and how contextual factors (local, national and/or international) have influenced evaluation design.
 - The purpose, scope (including the limits), and objectives of the evaluation, including primary and secondary audience/users are clearly explained and identified.
 - The evaluation processes and who were involved /consulted at each step were clearly explained.

3. **Evaluation Framework and Methodology** - The evaluation framework and methodology are specific and tailored to the evaluation context, and sufficient to address the purpose, scope, and objectives of the evaluation.
 - Relevant Adaptation Fund evaluation criteria have been selected, justified, and applied as per the definitions provided in the Evaluation Policy and accompanying guidance notes. Where there were excluded criteria, AF-TERG clearance has been secured.
 - Evaluation questions (and sub-questions where need) were clearly identified, correspond correctly with the selected Adaptation Fund evaluation criteria, and sufficient to address the evaluation objectives.
 - The methodology provides for multiple lines of inquiry/complementation of qualitative and quantitative techniques, and triangulation of data. If not, a clear justification for doing otherwise has been provided.
 - The methodology was applied as described.
 - Methodological limitations are acknowledged and their impact on evaluation design discussed. Limitations are acceptable and the risks they presented to the evaluation were sufficiently addressed.

4. **Methods and Data** - The methods are appropriate and can adequately address the evaluation questions. Data sources are appropriate, adequate, and sufficiently robust.
- Methods and how they were applied were described systematically and are appropriate for addressing the evaluation questions.
 - Where samples are needed, the sampling strategy is described and sample sizes are adequate.
 - Primary and secondary data sources used were appropriate, adequate, and reliable.
 - Different stakeholders at all levels were reached and consulted. If not, a justification for this was provided, and its impact on the quality of evidence and the evaluation was explained transparently.
 - If mixed methods were employed, their relationship and relative strengths are explained.
 - The methods enabled the collection and analysis of disaggregated data to show difference between groups (particularly women and girls, young people, and those in lower income quintiles).
5. **Analysis and Findings** - The findings and final ratings were well substantiated.
- The evaluation clearly indicate judgement criteria for assessing the evaluation questions and evaluation criteria, including for assessing the quality of project M&E and the likelihood that the achieved and expected outcomes (including in relation to climate change adaptation and resilience) would be sustained in the portfolio of completed projects.
 - The rubrics rating scale was applied properly to assess performance against the evaluation criteria.
 - The findings respond directly to the evaluation questions.
 - The evaluation report assessed the strength of evidence for each evaluation question and was clear on the weaknesses in evidence.
 - The evaluation analysed potential unanticipated or negative consequences.
 - The analytical process and specific analytical methods were specified and clearly applied.
 - Data sources substantiating findings were comprehensively identified.
 - There was a protocol or a process for handling competing interests, differences of opinion, disputes, and grievances. Was it described transparently in the report?
6. **Conclusions, Lessons, and Recommendations** - The conclusions, lessons, and recommendations are logical, coherent, well-substantiated, and practical/realistic. Forward looking lessons on climate change adaptation programming, and beyond the project boundaries are captured.
- Final evaluation findings and conclusions can inform strategic decision-making, demonstrate the relevance, effectiveness and efficiency of project design, objectives, and performance.
 - There is a logical linkage between conclusions, lessons, and recommendations.
 - The evaluation report clearly substantiated conclusions and recommendations with evidence.
 - Lessons identified can help to improve and instruct future climate change adaptation interventions, and specifically speak to issues related to maladaptation, CCA finance, and other key issues in CCA.

- Recommendations are specific, realistic for implementation and credible given time, resources, and capacities for implementation.
- Recommendations have been prioritized and ownership has been identified.

7. **Inclusion, Ethics, and Independence** - The evaluation processes were undertaken in an inclusive, ethical, and independent manner. The evaluation report represented the views of a diverse range of stakeholders.

- Different groups of stakeholders and beneficiaries such as women, children, indigenous people, etc. were meaningfully involved in the evaluation processes (i.e., planning, data collection, analysis, formulation of recommendations, etc.).
- The evaluation adequately assessed equitability and gender sensitive inclusivity in both the content and process aspects of the evaluation.
- Beneficiary and activity level data disaggregated by social criteria (e.g. sex, ethnicity, age disability, location, income, or education) to account for potential discrimination and exclusions.
- The evaluation upheld the Fund principles and priorities embodied in the Gender Policy and Action Plan of the Adaptation Fund.
- The evaluation report explained the ethical considerations within the evaluation, and how the evaluators adhered to ethical standards during the design and implementation of the evaluation.
- The report demonstrated the professional and cultural competencies of the evaluators (i.e., was there a meaningful mix of international, national, male, female, etc. evaluators).
- The report confirmed that the evaluators have not been or expect to be in the near future directly involved in and responsible for the policy setting, design, or management of the evaluation subject.
- The evaluation report confirms how it ensured that the various aspects of its evaluation, such as design, framework, data collection, analysis, findings, conclusions, and recommendations are free from external, political, personal, and organizational influence (i.e. personal or professional threat) and bias.
- Any actual or potential conflict of interest affecting the evaluation team is disclosed and appropriate mitigation strategy is explained.

8. **Management and Governance** - The evaluation has clear management and governance arrangements.

- The evaluation management and governance arrangements are clearly described and promoted the independence and credibility of the evaluation.
- Accountabilities, responsibilities, and lines of communication within the evaluation team, and between the evaluation team and IE structures were clear.
- There were appropriate structures and timing for the quality assurance of the report.
- Risks and challenges to the evaluation processes and reporting were identified and adequately addressed.

9. **Utility** - The evaluation is designed to meet the information and decision-making needs of the intended users and other stakeholders.

- The potential users and stakeholders, and the ways in which the evaluation will be used, as well as how it will be disseminated and communicated to the audience, have been identified (i.e., communication and dissemination plan).

- It is clear how the intended audience were involved during validation and feedback processes.
- The evaluation report was submitted within nine months of project completion.
- An evaluation management response was used to support evaluation follow-up.

Annex 3 Intended outcomes of projects/programmes covered by this synthesis

Country	Project Outcomes
Mauritius	<ol style="list-style-type: none"> 1. Application of Adaptation Measures for Coastal Protection: Increased adaptive capacity within relevant development and natural resource sectors. 2. Early Warning System: Reduced exposure at national level to climate-related hazards and threats. 3. Training: Strengthened institutional capacity to reduce risks associated with climate-induced socioeconomic and environmental losses. 4. Policy Mainstreaming: Improved policies and regulations that promote and enforce resilience measures. 5. Knowledge Dissemination and Management: Effective capturing and dissemination of lessons from the applied activities in the programme.
Cuba	<ol style="list-style-type: none"> 1. Reduced exposure to climate-related hazards and threats. 2. Strengthened institutional capacity to reduce risks associated with climate induced socioeconomic and environmental losses. 3. Increased ecosystem resilience in response to climate change and variability induced stress. 4. Diversified and strengthened livelihoods and sources of income for vulnerable people in targeted. 5. Improved policies and regulations that promote and enforce resilience measures.
Mali	<ol style="list-style-type: none"> 1. The resilience to climate change of local water supply systems in the regions of Mopti and Timbuktu is increased. 2. The production of local livelihood systems such as agriculture, fisheries, livestock, and forestry in the context of climate change is increased. 3. The capacities of local institutions and communities to better adapt to climate change are strengthened.
Egypt	<ol style="list-style-type: none"> 1. Improved adaptive capacity of the southern region of the country in the face of anticipated climate-induced reduction in food security through asset creation, knowledge/technology transfer, and capacity/skills development. 2. Government more committed to investing in and sustaining climate risk reduction strategies and measures.
Ghana	<ol style="list-style-type: none"> 1. Improved basin level management and planning of water resources that take into account climate change impacts on surface and groundwater sources. 2. Climate resilient management of water resources by at least 30 communities in northern Ghana. 3. Enhanced diversification of livelihoods of 50 communities in northern Ghana.
Uzbekistan	<ol style="list-style-type: none"> 1. Institutional and technical capacity for drought management and early-warning developed. 2. Climate-resilient farming practices established on subsistence dekhans farms. 3. Landscape-level adaptation measures for soil conservation to improve climate resilience for over one million hectares (ha) of land. 4. Knowledge of climate-resilient agriculture and pastoral systems in arid lands generated and disseminated.

Country	Project Outcomes
South Africa	<ol style="list-style-type: none"> 1. Small grants support concrete adaptation measures that strengthen livelihood strategies, adaptive capacity, infrastructure, and assets in two district municipalities in SA. 2. SGR and associated institutions are empowered to identify response measures to climate-induced vulnerabilities and implement relevant cc adaptation projects. 3. A methodology for enhancing direct access to climate finance is developed, based on lessons learned, providing recommendations for scaling up and replicating in SA and beyond.
Madagascar	<ol style="list-style-type: none"> 1. Knowledge base on best practices for climate resilience in rice, based on existing local knowledge and international research. 2. Sustainable increase in rice yields (using MIRR). 3. Ecosystem services maintained. 4. Post-harvest losses reduced. 5. Technical norms and standards in rice cultivation reviewed and where necessary modified to take climate change into account.
Uruguay	<ol style="list-style-type: none"> 1. Increased resilience of smallholder beneficiaries to climate variability and drought, measured by increased availability of water and forage, conservation of native grassland biodiversity, improved animal performance indicators, low mortality rates by animal category and stability of livestock composition over time. 2. Local institutional networks at the LU level that manage climate risk, involving young people and managing operational instruments that respond in case of emergency, in close coordination with the Rural Development Roundtables, the Early Climate Warning Systems developed by the MGAP and the National Emergency System. 3. To have the capacities and methodologies for systematic monitoring of CC and variability and their impact on agriculture, as well as having a catalogue of best practices for reducing vulnerability and improving resilience, innovative tools, and lessons learned from systematized experiences, endorsed by all stakeholders with regard to CC adaptation and with special reference to droughts.
Cambodia	<ol style="list-style-type: none"> 1. Technical expertise and a local enabling framework for forest restoration and eco-agriculture interventions. 2. Multi-use forests established and maintained and agriculture practices diversified / intensified. 3. Multi-use forests established and maintained and agriculture practices diversified / intensified.
Rwanda	<ol style="list-style-type: none"> 1. Adaptation to climate change (rainfall intensity and duration) through integrated land and water management to support climate-resilient production and post-harvest systems. 2. Support for the transition from unsustainable settlement patterns and exploitative farming practices to sustainable, diversified livelihoods. 3. Capacity building of local institutions to plan and implement climate resilient land and water management regimes and scale up effective adaptation strategies at the national and local levels.