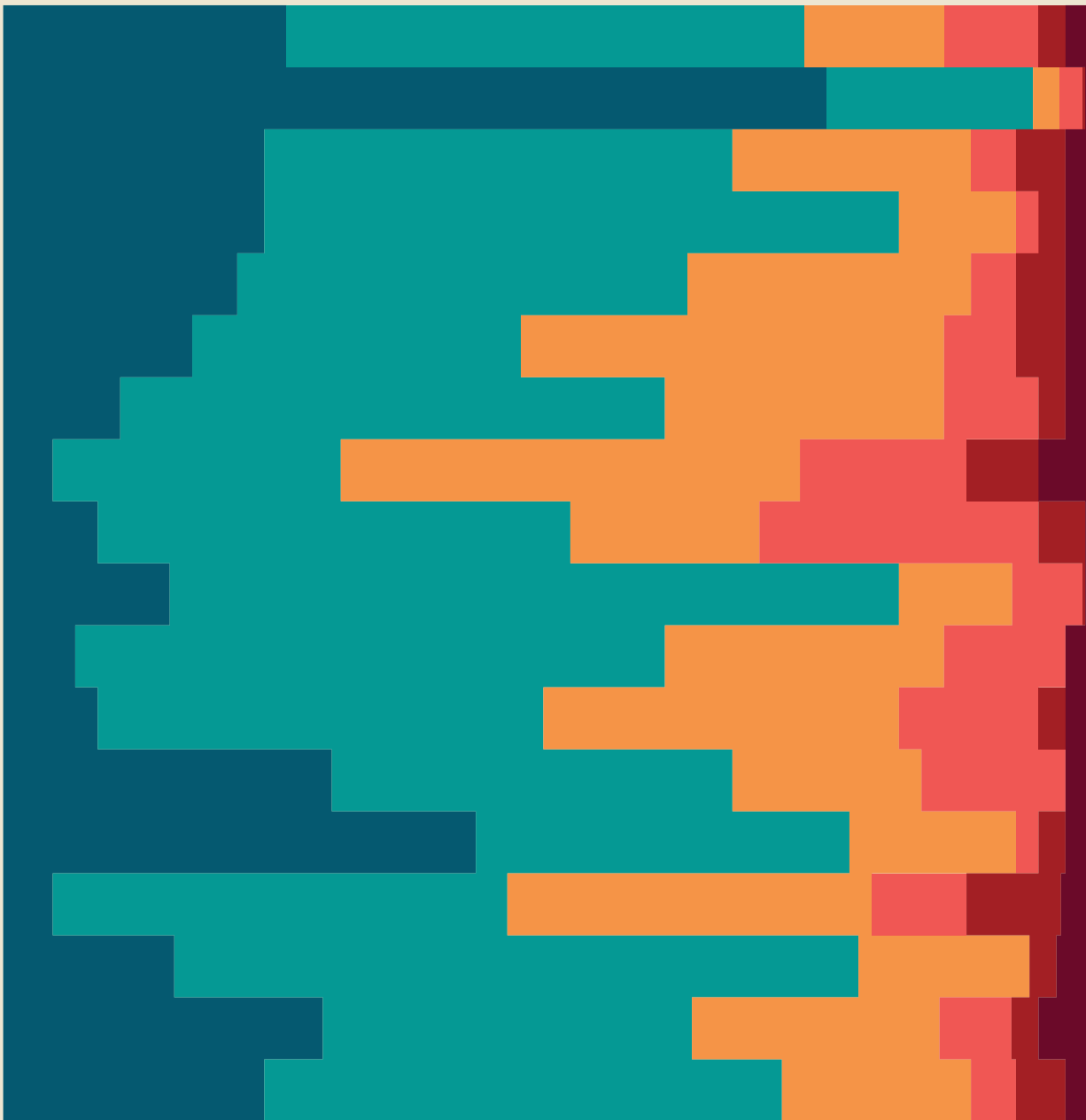

Evaluation Synthesis Report

2022 — 2023



Disclaimers

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory or city or area or its authorities, or concerning the delimitation of its frontiers or boundaries.

Mention of a commercial company or product in this document does not imply endorsement by the United Nations Environment Programme or the authors. The use of information from this document for publicity or advertising is not permitted. Trademark names and symbols are used in an editorial fashion with no intention on infringement of trademark or copyright laws.

This report is a product of the Evaluation Office of the UN Environment Programme (UNEP). The findings and conclusions expressed herein do not necessarily reflect the views of the UNEP Senior Management.

This publication may be reproduced in whole or in part and in any form for educational or non-profit services without special permission from the copyright holder, provided acknowledgement of the source is made. The United Nations Environment Programme would appreciate receiving a copy of any publication that uses this publication as a source. No use of this publication may be made for resale or any other commercial purpose whatsoever without prior permission in writing from the United Nations Environment Programme.

For further information on this report, please contact:

Evaluation Office of UNEP
P. O. Box 30552-00100 GPO
Nairobi Kenya
Tel: +254 20 762 3389
Email: Michael.Spilsbury@un.org

January 2024

All rights reserved.

© (2024) Evaluation Office of UNEP

Table of contents

Acknowledgements	vii	3 Analysis of project performance in 2022 and 2023	43
Abbreviations and Acronyms	viii	Overall Project Performance	47
Foreword	x	Overall project performance in management-led Terminal Reviews	48
Executive Summary	xii	Overall project performance presented by project size	48
1 Introduction	22	Quality of Project Design	49
Scope and objectives of the synthesis report	22	Strategic Relevance	50
Evaluation Office mandate	24	Effectiveness	51
Evaluation policy, roles and responsibilities	24	Availability of Outputs	52
Types of evaluation	25	Achievement of Project Outcomes	53
2 UNEP's evaluation criteria and coverage	27	Likelihood of Impact	54
Evaluation criteria and rating scales	27	Sustainability	55
Evaluation coverage	32	Socio-political Sustainability	57
Guidance and tools supporting evaluation and management-led Terminal Review processes	39	Institutional Sustainability	57
Learning from evaluations	40	Financial Sustainability	58
Limitations	41	Efficiency	58
		Financial Management	59

Monitoring and Reporting	61	Evaluation of the implementation of the Partnership for Action on Green Economy (PAGE) interagency programme 'Operational Strategy 2016-2020'	84
Monitoring Design and Budgeting	62		
Monitoring of Project Implementation	62	5 Compliance with evaluation recommendations	88
Project Reporting	63	Categorisation of recommendations	89
Factors affecting project performance	63	Compliance with recommendations	92
Preparation and Readiness	64	6 Special study – A synthesis of findings from assessments of project design quality	95
Quality of Project Management and Supervision	65	Background	95
Stakeholders' Participation and Cooperation	66	Analysis of project design quality assessments from project evaluations	97
Country Ownership and Driven-ness	67	Comparing project design quality ratings with evaluation project performance ratings	102
Communication and Public Awareness	68	Project Preparation (PDQ) vs Preparation and Readiness (TE)	102
Responsiveness to Human Rights and Gender Equality	69	Intended Results and Causality (PDQ) vs Effectiveness (TE)	104
Environmental and Social Safeguards	70	Sustainability / Replication and Catalytical Effects (PDQ) vs Sustainability (TE)	104
4 Key Findings from strategic evaluations	71	7 UNEP's updated evaluation policy and strategy	106
Evaluation of the UNEP subprogramme on Environmental Governance, 2014-2021	71	The role of the evaluation policy	107
Independent Evaluation of UNEP Policy and Strategy for Gender Equality and the Environment, 2015-2020	75	Five strategic and interrelated priority areas	107
Terminal Evaluation of the UNEP Inquiry into the design of a sustainable financial system and Interim Review of the UNEP/GEF Aligning the financial system and infrastructure and infrastructure investments with sustainable development (Components 1 and 2)	80	8 Performance of the Evaluation Office and evaluation quality	110
		The OIOS evaluation dashboard study of UNEP Evaluation Office	110

UN system-wide Action Plan on gender equality and the empowerment of women (UN-SWAP)	111	Annex 2 List of evaluations assessed for responsiveness to human rights and gender equality, 2022-2023	121
Annex 1 List of evaluations in the Evaluation Synthesis Report 2022-2023	115	Annex 3 UNEP's results definitions	125
		Annex 4 Weightings for evaluation criteria	132
List of figures			
Figure ES.1 Summary of project performance against evaluation criteria, 2022-2023	xiv	Figure 7. Alignment of evaluated projects to Sustainable Development Goals, 2022-2023	38
Figure ES.2 Summary of project performance against factors affecting performance, 2022-2023	xvi	Figure 8. Summary of project performance against evaluation criteria, 2022-2023	45
Figure ES.3 Overview of overall project performance, 2022-2023	xvii	Figure 9. Comparison of overall project performance, 2014-2023	47
Figure ES.4 Submission of evaluation management responses compliance levels, 2018-2023	xix	Figure 10. Overall project performance in management-led Terminal Reviews, 2022-2023	48
Figure ES.5 Recommendation compliance status, 2018-2023	xix	Figure 11. Overall project performance by project size, 2022-2023	49
Figure 1. Distribution of evaluated projects by subprogramme, 2022-2023	33	Figure 12. Quality of project design, 2018-2023	50
Figure 2. Distribution of evaluated projects by UNEP Division, 2022-2023	34	Figure 13. Comparison of strategic relevance (overall), 2014-2023	50
Figure 3. Distribution of evaluated projects by funding modality, 2022-2023	35	Figure 14. Effectiveness (overall), 2014-2023	51
Figure 4. Distribution of evaluated projects by geographic region, 2022-2023	36	Figure 15. Comparison of availability of outputs, 2014-2023	52
Figure 5. Evaluation consultants by gender, 2022-2023	36	Figure 16. Achievement of outcomes, 2014-2023	53
Figure 6. Evaluation consultants by geographic representation, 2022-2023	37	Figure 17. Likelihood of impact, 2014-2023	55
		Figure 18. Sustainability (overall), 2014-2023	56
		Figure 19. Project performance in socio-political, institutional, and financial sustainability, 2022-2023	57

Figure 20. Efficiency, 2014-2023	58	Figure 36. Compliance status of recommendation implementation plans closed, 2022-2023	93
Figure 21. Financial management (overall), 2014-2023	60	Figure 37. Comparison of recommendation compliance status by biennia, 2018-2023	94
Figure 22. Financial management by sub-criteria, 2022-2023	60	Figure 38. Range of ratings for project design quality assessments, 2017-2022	97
Figure 23. Monitoring and reporting, 2022-2023	61	Figure 39. Range of PDQ ratings, 2018-2022	98
Figure 24. Monitoring and reporting by sub-criteria, 2022-2023	62	Figure 40. Comparison of project design quality sub-criteria ratings	99
Figure 25. Projects rated 'Satisfactory' and better in the factors affecting performance, 2022-2023	63	Figure 41. Comparison of project design quality sub-criteria with 'Satisfactory' rating as the minimum acceptable standard	101
Figure 26. Preparation and readiness, 2014-2023	64		
Figure 27. Quality of project management and supervision, 2018-2023	65		
Figure 28. Stakeholders' participation and cooperation, 2014-2023	66	List of tables	
Figure 29. Country ownership and driven-ness, 2014-2023	67	Table 1. UNEP performance evaluation criteria and adjustments over time	28
Figure 30. Communication and public awareness, 2016-2023	68	Table 2. Performance rating scale	31
Figure 31. Responsiveness to human rights and gender equity, 2018-2023	69	Table 3. Number of completed evaluations by subprogramme and MTS/POW, 2010-2023	33
Figure 32. Environmental and social safeguards, 2020-2023	70	Table 4. Comparison of projects attaining S/HS ratings in the main evaluation criteria in 2020-2021 and 2022-2023	46
Figure 33. Distribution of the type of recommendations issued, 2022-2023	90	Table 5. Comparison of performance between availability of outputs and achievement of project outcomes, 2014-2023	54
Figure 34. Submission of management response compliance levels, 2018-2023	91	Table 6. Priority level of recommendations, 2022-2023	89
Figure 35. Overall acceptance of recommendations, 2018-2023	91	Table 7. UNEP results definitions relevant to evaluations	127
		Table 8. Weightings applied across evaluation criteria for overall performance ratings	132

Acknowledgements

The Evaluation Synthesis Report 2022-2023 was prepared by the Evaluation Office of UNEP under the overall guidance of Michael Spilsbury, Director, who also led content creation and preparation of Chapters 1, 2, 5-8. Janet Wildish, Senior Evaluation Officer, peer reviewed the report, while Pauline Marima, Evaluation Officer, led vital data analysis for Chapter 3, and Cecilia Morales, Evaluation Officer, prepared the section on UN-SWAP reporting for Chapter 8. Susanne Bech, Evaluation Officer, managed the overall production schedule and developed Chapter 4. We are grateful to Neeral Shah who conducted the detailed special study on the Quality of Project Design in Chapter 6.

The Evaluation Office would like to thank all the evaluators and evaluation specialists who shared their expertise over the biennium to the preparation of the evaluation reports mentioned in the Evaluation Synthesis Report.

Abbreviations and Acronyms

AF	Adaptation Fund	GPS	Gender Policy and Strategy
AfDB	African Development Bank	IFAD	International Fund for Agricultural Development
ANUBIS	Web-based project management and information system (GEF)	IFC	International Finance Corporation
CCA	Common Country Assessment	IGE	Inclusive Green Economy
COP	Conference of the Parties	InfoMEA	United Nations Information Portal on Multilateral Environmental Agreements
CPR	Committee of Permanent Representatives	ILO	International Labour Organization
EG	Environmental Governance	IPMR	Integrated Planning Management and Reporting System
e-PAS	Electronic Performance Appraisal System	KPI	Key Performance Indicator
EPI	Evaluation Performance Indicator	LDC	Least Developed Country
EU	European Union	MDB	Multilateral Development Bank
GCF	Green Climate Fund	MEA	Multilateral Environmental Agreement
GEEW	Gender Equality and Empowerment of Women	MOU	Memorandum of Understanding
GEF	Global Environment Facility	MPTF	Multi-Partner Trust Fund
GePA	Gender Plan of Action	MTS	Medium-Term Strategy
GEWE	Gender Equality and Women's Empowerment	OECD-DAC	Organisation for Economic Co-operation and Development's Development Assistance Committee
GFP	Gender Focal Points	OIOS	Office of Internal Oversight Services
GGEO	Global Gender and Environment Outlook	PAGE	Partnership for Action on Green Economy
GGKP	Green Growth Knowledge Platform	PIMS	Project Information Management System
GSU	Gender and Safeguards Unit		
GEF	Global Environment Facility		

PCAU	Programme Coherence / Assurance Unit	UNEA	United Nations Environment Assembly
PDQ	Project Design Quality	UNEG	United Nations Evaluation Group
PoW	Programme of Work	UNEP	United Nations Environment Programme
PPD	Policy and Programme Division	UNEP-FI	UNEP Finance Initiative
ProDoc	Project Document (UNEP)	UNFPA	United Nations Population Fund
PRC	Project Review Committee	UNICEF	United Nations Children's Fund
RCO	Resident Coordinator Office	UNIDO	United Nations Industrial Development Organization
SDG	Sustainable Development Goal	UNITAR	United Nations Institute for Training and Research
SIDA	Swedish International Development Cooperation Agency	UNOPS	United Nations Office for Project Services
SMT	Senior Management Team	UNSDCF	United Nations Sustainable Development Cooperation Framework
SP-EG	Subprogramme on Environmental Governance	UN-SWAP	United Nations system-wide Action Plan on Gender Equality and the Empowerment of Women
TE	Terminal Evaluation	USA	United States of America
ToC	Theory of Change	USD	United States Dollars
TR	Terminal Review	V20	Vulnerable Twenty Group
UN	United Nations	WMO	World Meteorological Organization
UNCT	United Nations Country Team	WTO	World Trade Organization
UNDA	United Nations Development Account		
UNDP	United Nations Development Programme		

Foreword

The Evaluation Synthesis Report 2022-2023 summarises findings derived from the work of UNEP's Evaluation Office in 2022 and 2023. The biennial programme of evaluations includes within its focus: sub-programme and policy level evaluations; evaluations of completed projects; validation of project level management-led Terminal Reviews; the development of performance assessment methods; reporting and communicating findings and results; and evaluation follow-up and recommendation compliance. All these activities are undertaken in a fully independent manner. The Evaluation Office regularly engages with the Executive Director and the Senior Management Team with evaluation being a regular part of UNEP's Quarterly Business Review.

The report highlights the performance trends across 46 UNEP projects that were evaluated during the period and sets them in the context of project performance trends across five biennia. A positive trend was evident with an increasing proportion of projects rated in the top two performance categories for their overall performance. The report presents a number of important lessons derived from project implementation experiences that should make interesting reading for all UNEP Project Managers.

However, the report highlights some areas for continued management attention. The long-term impact for completed projects remains somewhat uncertain suggesting that UNEP project designs and implementation actions should focus on developing more robust strategies to sustain project outcomes and place more emphasis in terms of resources and management actions on the processes leading from outcomes towards longer term effects.

Follow up to evaluation is improving. The report highlights improved compliance with evaluation recommendations compared to previous reporting periods, and the rate of completion of evaluation implementation plans has, for the first time, reached 100%.

In 2023, the Evaluation Office was, itself, the subject of a positive independent performance assessment conducted by the OIOS and has regularly been called upon as a source of knowledge and advice by the Business Transformation and Accountability Division on the UN Secretariat. Over many years, positive external assessments of the UNEP Evaluation Office have demonstrated the

independence and credibility of the function and bolstered the confidence that the organisation can place in its findings.

In the next biennium, UNEP can look forward to receiving further evaluative findings that can provide valuable insights to design, implementation and management challenges at organisation, policy, programme and project levels.

Michael J Spilsbury

Director

UNEP Evaluation Office

Executive summary

The mandate for coordinating, conducting and overseeing evaluation in UNEP is vested in the Evaluation Office. This mandate covers all programmes and projects of the Environment Fund, related trust funds, earmarked contributions and projects implemented by UNEP under the Global Environment Facility (GEF). The Office undertakes a variety of evaluations and management studies, in accordance with the requirements of the United Nations General Assembly, the United Nations Environment Assembly of UNEP, and in accordance with the Norms and Standards for evaluation of the United Nations system.

The work of the Evaluation Office enhances accountability, transparency and learning. Evaluations generate evidence to identify ‘what works’ and ‘what doesn’t’ and provide feedback for the improvement of planning and management processes. Project evaluations undertaken by the Evaluation Office respond to implementation timelines; evaluations are scheduled and initiated as projects reach, or approach, their operational completion.

The Evaluation Synthesis Report 2022-2023 serves as part of the Evaluation Office’s mission to promote a results-oriented approach within UNEP, emphasizing organizational learning, informed decision-making, and accountability. Aligned with the Secretary General’s guidelines, the report aims to systematically assess the relevance, efficiency, effectiveness, and impact of UNEP’s activities. It provides an evaluative assessment of UNEP’s programme and project performance in 2022 and 2023, a period marking the beginning of the UNEP Medium-Term Strategy (MTS) 2022-2025. The document provides an overview of the evaluation criteria and coverage employed by the UNEP Evaluation Office. It outlines nine evaluation criteria, each with specific sub-criteria, used to assess UNEP’s project performance. The criteria cover aspects such as strategic relevance, project design, external context, effectiveness, financial management, efficiency, monitoring and reporting, sustainability, and factors affecting performance and cross-cutting issues.

This Evaluation Synthesis Report draws on in-depth evaluations of various UNEP initiatives and projects, offering an analysis of 46 evaluations conducted in 2022 and 2023, and examining performance trends over four previous biennia. The primary goal is to help UNEP reflect on and learn from its programme performance through evaluative evidence and lessons. The report

also presents highlights drawn from a selection of evaluations of high strategic importance including: the UNEP subprogramme on Environment Governance; an evaluation of the implementation of Policy and Strategy for Gender Equality and the Environment (2015-2020); a terminal evaluation of the UNEP Inquiry into the Design of a Sustainable Financial System; and the evaluation of the Partnership for Action on Green Economy (PAGE) Interagency Programme's Operational Strategy (2016-2020).

Coverage and performance of projects completed in 2022 and 2023

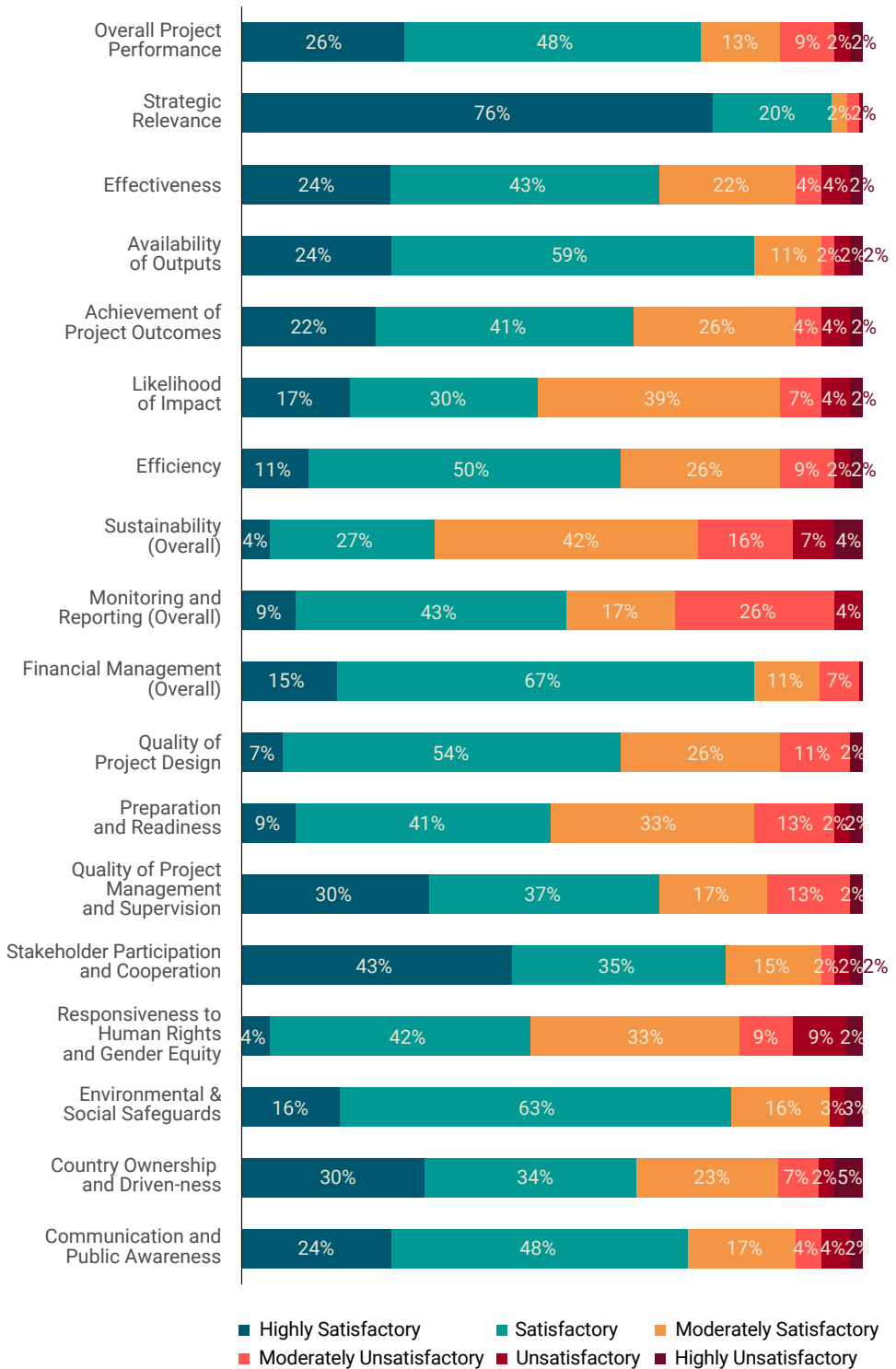
The report highlights the number of evaluations conducted across different subprogrammes and time periods, emphasizing the efforts to maximize coverage while maintaining quality. It further explores the distribution of evaluated projects by UNEP division, funding modality, geographic region, and Sustainable Development Goals (SDGs). The challenges and efforts to enhance evaluation coverage, representation, and diversity among evaluation consultants are also discussed.

All projects are evaluated against a standard set of evaluation criteria that are consistent with international good practice. Performance against all criteria is rated on a six-point scale from 'Highly Unsatisfactory' through to 'Highly Satisfactory'. The benchmark for good performance is set by the UNEP Evaluation Office at ratings of 'Satisfactory' or 'Highly Satisfactory' (S/HS).

Individual evaluation criteria are aggregated in each evaluation using a 'weighted' ratings scale in order to derive an overall performance rating. Within this scale the greatest emphasis is placed on the Achievement of Outcomes (30% of the overall performance score) and the Sustainability/durability of outcomes (20% of the overall performance score). This weighting reflects UNEP's emphasis on assessing performance from a results (rather than activities) perspective. **Figure ES.1** below shows the performance of projects evaluated in 2022 and 2023 against the standard evaluation criteria.

A positive performance trend continues for the proportion of projects attaining ratings in the Satisfactory range (HS, S and MS) – this remains high across most evaluation criteria as it has for the last five biennia. Across the same 10-year period, the 2022-23 biennium shows the highest proportion of projects rated 'Highly Satisfactory' for their overall performance, with 26%. The average proportion of projects attaining an S/HS rating for overall project performance across the four previous biennia was 56% compared to 74% of projects in 2022-2023.

Figure ES.1 Summary of project performance against evaluation criteria, 2022-2023



Another noteworthy increase was in the proportion of projects attaining S/HS ratings across the 2022 and 2023 sample for the criterion on Financial Management (87%); an increase of 41 percentage points when compared to the sample for the previous biennium (46% in 2020-21). Likewise, there have been increases in the sample for 2022 and 2023 in performance against the following criteria: Effectiveness (28 percentage point increase); and Sustainability (21 percentage point increase).

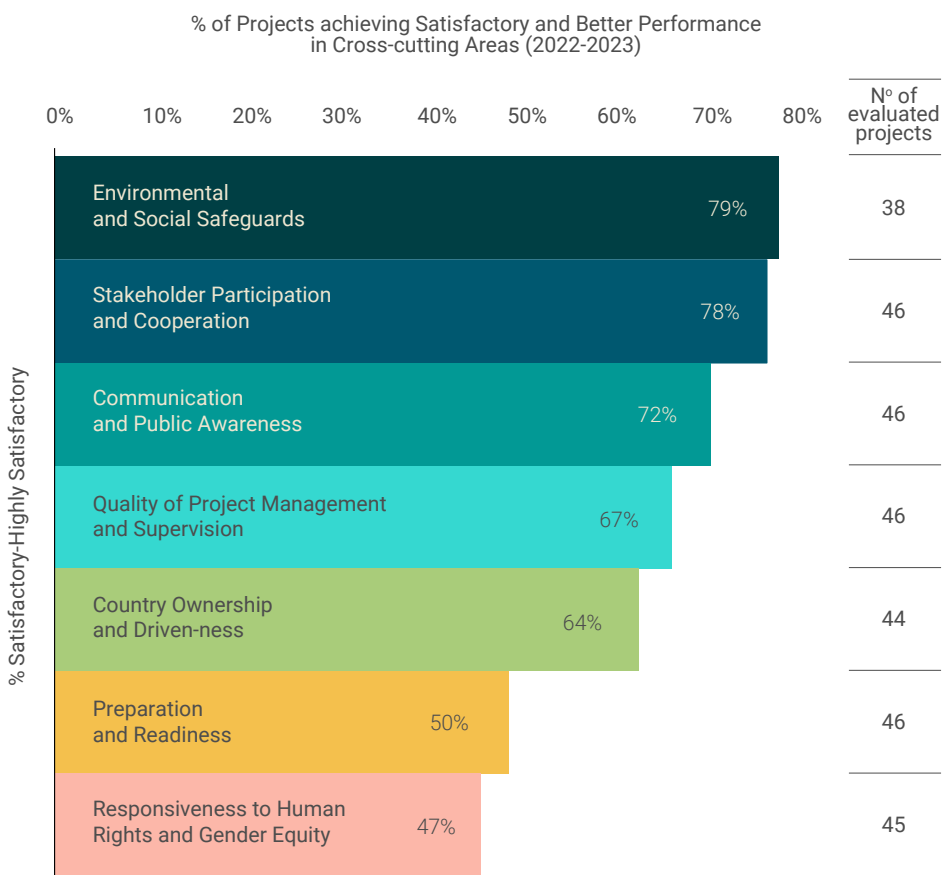
The evaluation of the Effectiveness criterion requires the aggregation of three evaluation sub-criteria: Availability of Outputs, Achievement of Outcomes and Likelihood of Impact. The percentage of projects attaining ratings of S/HS for the Availability of Outputs was 83% whilst the percentages for latter two evaluation sub-criteria were 63% and 47% respectively. The corresponding percentages in the previous 2020-21 biennium were 71%, 34% and 34% respectively. The Evaluation Office places a high weighting on the Achievement of Outcomes sub-criterion. Whilst performance improvements were recorded in the sample for 2022 and 2023 for this criteria, longer-term performance trends suggest a need for continued attention to project design and implementation actions that place a strong emphasis, in terms of resources allocated and management attention, on activities that drive the intended change processes beyond output delivery and on towards outcomes and impacts.

Approximately half of the projects evaluated in the last two biennia have not reached the desired performance target (HS/S) for the Monitoring and Reporting criterion. One of the key areas of weakness that continues to be observed, especially in projects not attaining the S/HS level, is the lack of distinction made between reporting on project implementation (what has happened) and monitoring implementation against the expected results. Learning from past evaluations also emphasize the need to strengthen the institutionalization of project documentation to ensure that critical project information is secured with UNEP regardless of staff turnover or operational completion of the project. Since monitoring and reporting are aspects of project management that are fully under UNEP's control or oversight, this is an area for continued management attention.

In an evaluation, sustainability is understood as the probability of project outcomes being maintained i.e. their endurance/persistence beyond project completion. In 2022 and 2023, 31% of the projects evaluated attained an S/HS rating for Sustainability 're-bounding' from the low of 10% observed in the sample of projects evaluated in the previous biennium. Nevertheless, the need to pay particular attention to the development of exit strategies and on creating the conditions that help to sustain direct outcomes in both project design processes and during project implementation remains. Strategies that enhance country ownership can also help ensure the continuance of project outcomes.

The report also discusses factors influencing project performance. Six key factors are analysed: Preparation and Readiness, Quality of Project Management and Supervision, Stakeholders’ Participation and Cooperation, Country Ownership and Driven-ness, Communication and Public Awareness, and Responsiveness to Human Rights and Gender Equality. **Figure ES.2** below shows performance against these factors in 2022 and 2023.¹

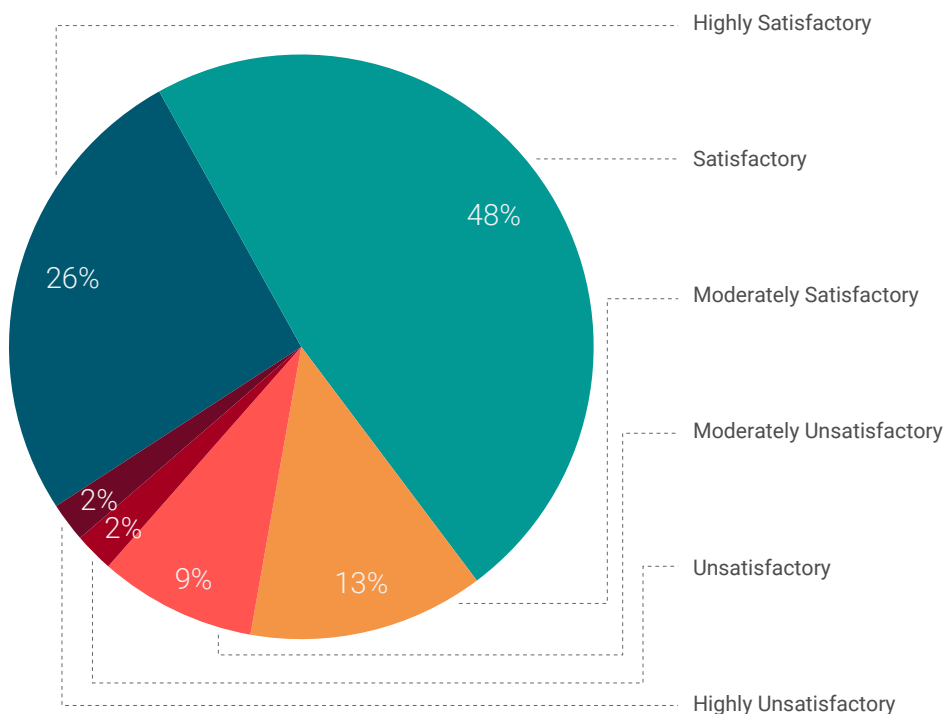
Figure ES.2 Summary of project performance against factors affecting performance, 2022-2023



In summary, the analysis provides a comprehensive overview of project performance, highlighting the positive patterns evident in the sample for 2022 and 2023 whilst highlighting areas for further improvement. The performance recorded against most of the criteria improved as compared to the previous biennium. There were no criteria where performance was substantially lower than the 2020-21 biennium. **Figure ES.3** below shows overall project performance.

¹The number of evaluated projects with a rating of factors affecting performance varies as some evaluations do not apply all evaluation criteria, for example, subprogramme evaluations.

Figure ES.3 Overview of overall project performance, 2022-2023



Biennium		2022-2023
Number of evaluated projects		46
% of projects with overall performance rated	Highly Satisfactory	26%
% of projects with overall performance rated	Satisfactory or better (i.e., S / HS)	74%
% of projects with overall performance rated	Moderately Satisfactory or better	87%

Top rated evaluation criteria in 2022-23
Percentage of projects rated against these criteria achieving 'Satisfactory' or better ratings (H/HS) is >70%

- 96%** Strategic relevance
- 83%** Availability of Outputs
- 82%** Financial management
- 78%** Stakeholder participation and cooperation
- 72%** Communication and public awareness

Areas for improvement
Percentage of projects rated against these criteria with ratings in the 'Unsatisfactory' range (MU-HU) is >15%

- 30%** Monitoring and reporting
- 27%** Sustainability of Project Outcomes
- 20%** Responsiveness to gender and human rights
- 17%** Preparation and readiness
- 15%** Quality of project management and supervision

In 2022, UNEP approved a new Evaluation Policy which introduced a new performance assessment modality, namely management-led Terminal Reviews. In 2022 and 2023, a total of 15 management-led Terminal Review reports were completed and submitted to the Evaluation Office for independent validation and an assessment of the report's quality. These projects were rated against the same performance criteria as used in independent terminal evaluations. The data show that the majority of these projects were rated at the 'Satisfactory' level for overall performance by both the Reviewer and the Evaluation Office (71% and 73% respectively). The small sample size limits the inferences that can be drawn from this sample.

Value of evaluated projects

The sum total of the reported expenditures declared by the 46 projects evaluated in this biennium was > **USD 157 million**. This amount included GEF /Adaptation fund grants, extra-budgetary funding, Environment Fund contributions, co-financing (cash) and resources covered by strategic cooperation agreements (e.g., European Commission). This amount did not include sources that were recorded as 'in-kind' co-financing. Reporting on project co-financing was not always available, accurate or consistent, and in some cases, there is no disaggregation between 'cash' and 'in-kind' financing, making it difficult to determine the true value of the actual co-financing investment. The value of the 15 projects that met their performance assessment requirements through management-led Terminal Reviews and were independently validated by the Evaluation Office totalled **USD 52,506,910**.

Evaluation recommendation compliance

Each evaluation process managed by the Evaluation Office is followed by a recommendation compliance process. In 2022 and 2023 there has been continued improvement in the submission of management responses to evaluations in the form of recommendation implementation plans. **Figure ES.4** shows that in 2022 and 2023, the percentage of evaluations that did not receive a management response was zero, a marked improvement from the 10%, and the 25% reported in the previous two biennia. However, of the 34 recommendation implementation plans submitted only 39% (13) were submitted within the required² timeframe compared to 45% in the last biennium. The preparation of a management response in the form of a recommendation implementation plan is a mandatory requirement for all evaluations. Compliance rates for completion of management responses to evaluation have now reached 100% although the timelines for these standard requirements need to be respected.

² Completed Recommendation Implementation Plans are required to be submitted to the Evaluation Office no more than one month after being shared with the relevant UNEP Project Manager.

Improvements in recommendation implementation compliance are evident in 2022 and 2023. As **Figure ES.5** shows, a larger proportion (59%) of recommendations were closed as 'Compliant' in 2022 and 2023 compared to the previous biennium (53%). The proportion of recommendations closed as 'Not compliant' also fell from 10% in the previous biennium to 2% in 2022 and 2023. Overall, there has been improvement in recommendation compliance rates. The Evaluation Office attributes much of the improvement to the use of corporate dashboards that are reviewed by the Senior Management Team every quarter. Continuous management attention is recommended for sustained progress in meeting UNEP's performance targets.

Figure ES.4 Submission of evaluation management responses compliance levels, 2018-2023

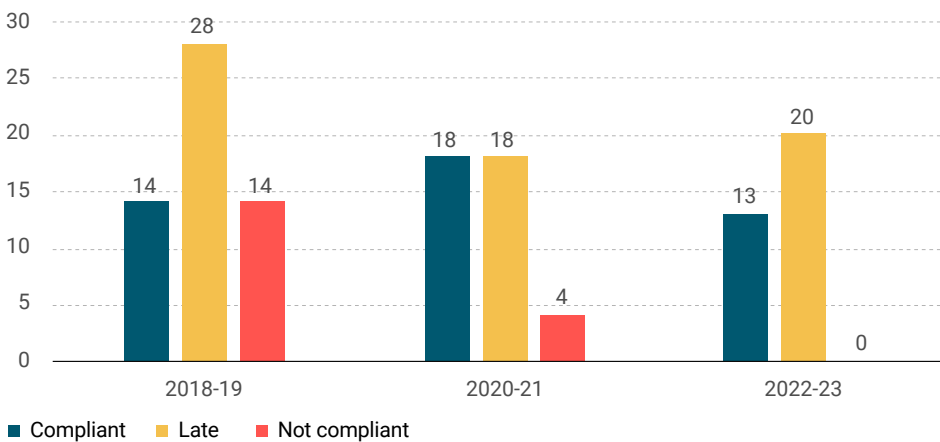
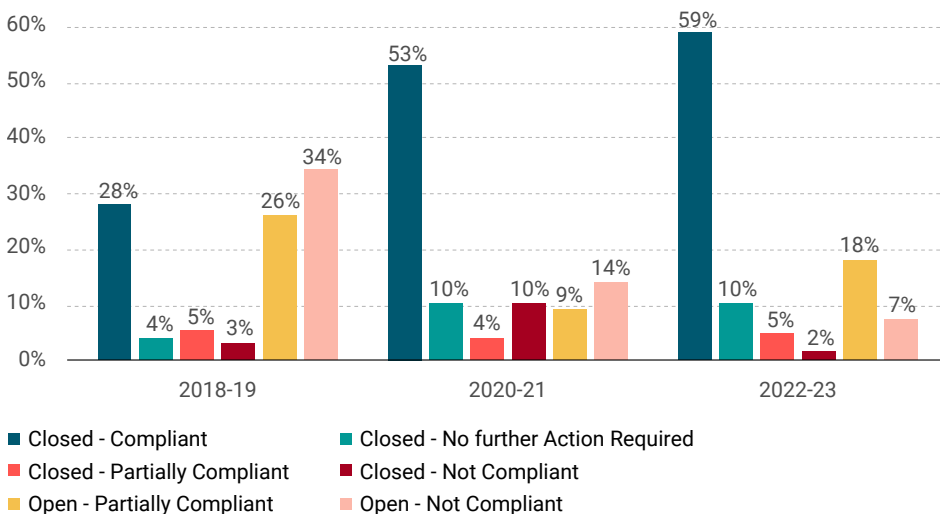


Figure ES.5 Recommendation compliance status, 2018-2023



Special Study on Project Design Quality

This chapter of the report provides a synthesis of findings from assessments of project design quality conducted by the Evaluation Office based on data collected between 2017 and 2021. It examines trends and patterns in project design quality and explores the relationship between project design assessments and subsequent project performance evaluations. The study found that over time, there is a trend indicating improved project design quality in projects evaluated between 2018 to 2021. Weaknesses in project design aspects identified included ‘Intended Results and Causality,’ ‘Sustainability/Replication and Catalytic Effects,’ ‘Risk Identification and Social Safeguards,’ and ‘Learning, Communication, and Outreach.’ It found that the role of the Project Review Committee in improving project design quality is not fully utilized because some identified weaknesses are not subsequently addressed by project teams. The study offers insights into areas for potential improvement in project design, approval processes, and implementation within UNEP.

Performance of the UNEP Evaluation Office

In May 2023, the United Nations Office of Internal Oversight Services (OIOS) published an assessment of UNEP’s Evaluation Office through its ‘Evaluation Dashboard’ report, acknowledging that UNEP possesses a robust evaluation system. The report highlighted improvements from previous assessments in various areas, including evaluation planning, report quality (81% rated ‘Good’ or ‘Very Good’, up from 64%), and the quality of recommendations (76% rated ‘Good’ or ‘Very Good’, up from 45%). Although UNEP’s reports meeting UN-SWAP requirements increased to 23% (up from 18%), this area still requires improvement, (as emphasised by the 2023 UN-SWAP assessment of ‘approaches requirements’). Notably, the UNEP Evaluation Office was recognized as the most productive within the UN Secretariat, producing 41 evaluation reports in 2020 and 2021. Additionally, OIOS commended UNEP for preparing high-quality evaluation reports demonstrating clear linkages between findings, conclusions, and recommendations, including in-depth analyses of gender considerations, human rights, and environmental issues. OIOS’s overall assessment affirmed the strength and effectiveness of UNEP’s evaluation practices.

New evaluation policy, operational strategy and manual

UNEP updated its Evaluation Policy in 2022, to be aligned with the Medium-Term Strategy 2022-2025, the policy emphasizes transformative actions and the centrality of science in decision-making. The revised policy integrates UN reform efforts to strengthen evaluation capacity and enhance program planning and reporting. UNEP’s Evaluation Office introduced a strategic approach into

the new policy and developed an Evaluation Operational Strategy focused on five priorities: developing a strategic evaluation agenda, enhancing the use of evaluation, maintaining quality in evaluation processes and products, building capacity for project-level assessments, and establishing a stable resource base. Key Performance Indicators are defined for each priority, reflecting UNEP's commitment to organizational learning, evidence-based decision-making, and transparency. The new Evaluation Policy and Operational Strategy are supported by a new Evaluation Manual.

Chapter 1

Introduction

Scope and objectives of the synthesis report

This evaluation synthesis report has been prepared as part of the mission of the Evaluation Office to promote a results-focus in UNEP that reflects an organizational culture of learning, informed decision-making and accountability. According to the Secretary General's bulletin on programme planning, monitoring and implementation (ST/SGB/2018/3), which consolidates the General Assembly decisions on the evaluation function, *"The objective of evaluation is: (a) To determine as systematically and objectively as possible the relevance, efficiency, effectiveness and impact of the Organization's activities in relation to their objectives; (b) To enable the Secretariat and Member States to engage in systematic reflection, with a view to increasing the effectiveness of the main programmes of the Organization by altering their content and, if necessary, reviewing their objectives."*

This report is prepared as an inter-sessional document of the UN Environment Assembly of UNEP and serves as part of the input of UNEP to the Secretary-General's report on evaluation to the General Assembly. The report provides stakeholders such as Governments, UNEP senior management and partners with an evaluative assessment of UNEP's programme and project performance in 2022 and 2023.

The biennium marked the beginning of the UNEP Medium-Term Strategy (MTS) 2022-2025, which rests on the recognition of the existence of three planetary crises – climate change, biodiversity loss and pollution – which put us at risk of irreversibly changing our relationship with the natural world. The MTS outlines a set of transformative shifts that target the drivers of climate change, biodiversity loss and pollution, and looks at their impact. It maps out the actions needed to reshape our consumption and production patterns towards sustainability, framing the contribution of UNEP in the context of the 2030 Agenda for Sustainable Development and the decade of action to deliver the Sustainable Development Goals and beyond, and in synergy with multilateral environmental agreements and other internationally agreed frameworks as well as leveraging the wider United Nations system for stronger, more coordinated and mutually supportive environmental action. Institutionally, the MTS meant that the previous format of subprogrammes was re-channelled towards three thematic subprogrammes (climate stability, living in harmony with nature and towards a pollution-free planet) supported by two foundational subprogrammes (science-policy and environmental governance) and two enabling subprogrammes (financial and economic transformations and digital transformations). Further in 2022, a 'standardized delivery model' (or New Delivery Model) was designed to operationalise the MTS, and the subprogramme structure.

The report utilizes information drawn from in-depth evaluations including: an evaluation of the Environmental Governance Sub-Programme; an independent evaluation of UNEP's Policy and Strategy for Gender Equality and the Environment; an evaluation of the UNEP Inquiry into the Design of a Sustainable Financial System and components of the UNEP/GEF project Aligning the Financial System and Infrastructure and Infrastructure Investments with Sustainable Development; and an evaluation of the Implementation of the Partnership for Action on Green Economy (PAGE) Interagency Programme 'Operational Strategy' and 46 other evaluations of projects contributing to the UNEP Programme of Work. The report presents **an analysis of the 46 evaluations conducted in 2022 and 2023** and includes projects and interventions implemented over the period from 2009 to 2023. The majority of projects evaluated (82%) began implementation between 2013 and 2017. The Evaluation Synthesis Report also contains a review of the implementation and compliance status of evaluation recommendations and presents external assessments of the performance of the Evaluation Office.

The main objective of the report is to help UNEP reflect on, and learn from, its programme performance through evaluative evidence and lessons drawn from multiple evaluation processes.

Evaluation Office mandate

The Evaluation Office conducts various types of evaluations and management studies in accordance with the requirements of the United Nations General Assembly, the United Nations Environment Assembly and the Norms and Standards for Evaluation of the United Nations system.

The mandate for evaluations in UNEP covers all programmes and projects of the Environment Fund, related trust funds, earmarked contributions and projects implemented by UNEP under the Global Environment Facility (GEF), the Green Climate Fund (GCF) and under partnership agreements. The mandate for conducting evaluations derives from UNEP Governing Council decisions, the UN Secretariat's Administrative Instruction on Evaluation ST/AI/2021/3 and several UN General Assembly Resolutions, summarized in the regulations and "Rules Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods of Evaluation" (ST/SGB/2018/3).

Evaluations are conducted in an independent manner and the Evaluation Office reports on evaluation findings without interference. To give it independence from substantive Divisions and operational sub-programmes, the Evaluation Office Director reports to the Executive Director, and through the Executive Director, to the Governing Bodies - Committee of Permanent Representatives (CPR) and United Nations Environment Assembly (UNEA). UNEA / CPR reviews the Evaluation Synthesis Report, which encompasses findings derived from all UNEP evaluations. Summaries of completed evaluations are shared in UNEP's quarterly reporting to the CPR.

Evaluation policy, roles and responsibilities

The UNEP Evaluation Policy guides the work of the Evaluation Office. It explains the objectives, roles and functions of evaluation within UNEP, the institutional framework and processes by which it is operationalized. The policy seeks to increase transparency, coherence and efficiency in generating and using evaluative knowledge for organizational learning and effective management for results, and to support accountability. To comply with new UN Secretariat requirements (ST/AI/2021/3), and respond to the UNEP MTS 2022-2025, a new Evaluation Policy was prepared during 2021, and approved in October 2022 (see **Chapter 7**).

The Executive Director of UNEP is guardian of the evaluation function and responsible for ensuring that the evaluation policy is implemented on behalf of the Secretary-General and Member States, overseeing the overall evaluation function, and ensuring the function is adequately resourced.

The Evaluation Office reporting line is directly to the Executive Director and the office is responsible for implementation of the evaluation work plan by conducting and managing independent evaluations at Medium-Term Strategy MTS/ Programme of Work (PoW), subprogramme, portfolio and project levels. It ensures quality in evaluations conducted, provides analysis of findings and lessons learned, prepares the Evaluation Synthesis Report and disseminates evaluation findings and results to UNEP, Member States and stakeholders. In addition, the Evaluation Office promotes the uptake of lessons and tracks compliance with evaluation recommendations.

UNEP Senior Management discusses and comments on strategic evaluations, ensures management responses to all evaluations are provided and that findings are incorporated in the design and implementation of all programme activities. Senior managers are also invited to provide input into the evaluation work plan. Other key responsibilities held by Technical Divisions, which are directly accountable for project implementation, progress monitoring and reporting, are to ensure that accepted evaluation recommendations are acted upon and that evaluation findings inform strategic planning processes. Project management staff and Subprogramme Coordinators identify projects reaching operational completion, coordinate their review of draft evaluation reports and prepare management responses.

Types of evaluation

Types of evaluation take four main forms: i) strategic and cross-cutting thematic evaluations including subprogramme evaluations; ii) impact/influence evaluations; iii) programme/ portfolio or project evaluations / validations; and iv) joint evaluations with other United Nations agencies, donors and partners.

The selection of evaluations is driven by the need to evaluate the performance of the Programme of Work and represent its key features in terms of; themes, subprogrammes, operational divisions/offices, geographic distribution of efforts and funding sources. As such, all UNEP subprogrammes are, since ST/AI/2021/3 was issued in late 2021, to be evaluated across a six-year period and a sample of projects reaching operational completion are to be selected by the

Evaluation Office for independent evaluation. The factors that inform purposeful selection of projects for the sample are set out in the new Evaluation Manual³.

The Evaluation Office closely follows-up on the implementation of all accepted evaluation recommendations and provides technical backstopping to project and programme managers undertaking management-led Terminal Reviews by: providing a suite of tools, templates and guidelines to support the project review process; providing technical guidance on a needs-basis, including guidance on consultant suitability; and completing a quality assessment and validation of the final project Terminal Review report.

UNEP staff responsible for projects that are not selected for independent evaluation are required to prepare a management-led Terminal Review if the project exceeds the financial expenditure threshold specified in the Evaluation Manual.

Guidelines, formal requirements and practical advice in planning for evaluations and management-led reviews is specified in detail in the UNEP Programme and Project Management Manual and further elaborated on the UNEP Evaluation website⁴.

³ <https://wedocs.unep.org/20.500.11822/42025>

⁴ www.unep.org/evaluation-office

Chapter 2

UNEP's Evaluation Criteria and Coverage

Evaluation criteria and rating scales

The UNEP Evaluation Office assesses performance against a total of nine evaluation criteria, six of which have sub-criteria (see **Table 1** below). While these criteria reflect standard internationally accepted⁵ performance criteria (Relevance, Efficiency, Effectiveness, Impact and Sustainability), they are also adapted to be relevant to those areas of performance of most interest to UNEP. For example, UNEP's evaluation criteria allow for assessments against 'financial management' and 'monitoring and reporting', but also cover a range of other factors affecting performance and cross-cutting issues.

⁵ The evaluation criteria described by the OECD/DAC are widely accepted as an international standard. UNEP Evaluation Office also works in accordance with the descriptions reflected in the UN Evaluation Group Norms and Standards.

This set of performance evaluation criteria have evolved over time to: a) be responsive to UNEP’s policies and areas of priority; b) make necessary areas of learning visible and c) create a comprehensive and transparent framework against which all performance aspects of UNEP’s work can be evaluated. In **Table 1** below the major adjustments made in recent years are noted.

Table 1. UNEP performance evaluation criteria and adjustments over time

Evaluation criteria	
A. Strategic Relevance	
1. Alignment to the UNEP Medium Term Strategy (MTS), Programme of Work (POW) and Strategic Priorities	Prior to 2018 this criterion included a sub-category related to gender. At that time this placement reflected the place of gender in UNEP’s policies and strategic priorities. In 2018-19 ‘gender’ was moved to its own sub-category under cross-cutting issues to better reflect its central importance, along with human rights, during project implementation.
2. Alignment to Donor/GEF/ Partner strategic priorities	
3. Relevance to global, regional, sub-regional and national environmental priorities	Project alignment with the Sustainable Development Goals (SDGs) is covered under sub-criterion A.3.
4. Complementarity with relevant existing interventions/ coherence	The sub-criterion ‘complementarity’ (A.4) is consistent with the latest (2019) OECD criterion of ‘coherence’.
B. Quality of Project Design	
	Assessment against this dimension has been longstanding – with formal ratings for the criterion continuous since 2017. The assessment template used in evaluations, was the original basis for, and is consistent with that used during the project design approval process.
C. Nature of External Context	
	Introduced in 2018, this is not a performance criterion, and its rating is not used in the calculation of overall project performance. However, when a project is found to have faced some specific unfavourable external conditions that could not have been anticipated, the ratings on Effectiveness and Sustainability may be adjusted to take unexpected disruptions into account (e.g. conflict, natural disasters and political upheaval). Adjustments in the face of the COVID-19 pandemic are likely to be addressed here in the coming years.


Evaluation criteria	
D. Effectiveness	
1. Availability of outputs	Since 2019, the emphasis has been placed on the availability of Outputs to intended beneficiaries rather than the less specific term 'delivery' of Outputs. This is in keeping with UNEP's revised results definitions, 2019.
2. Achievement of project outcomes	Emphasis is placed on those outcomes (Project Outcomes) that a project is expected to achieve within its project timeframe and with its secured funding.
3. Likelihood of impact	Since 2018 an Excel-based tool that supports a combined assessment of a) the extent to which Outcomes have been achieved and b) the extent to which contributing conditions needed to create long-lasting change (i.e. assumptions and drivers) have been seen to hold, has been tested and strengthened.
E. Financial Management	
1. Adherence to UNEP's policies and procedures	'Assessment against this sub-criterion has been continuous since 2019
2. Completeness of project financial information	Assessment against this sub-criterion has been continuous since 2018.
3. Communication between finance and project management staff	Assessment against this sub-criterion has been continuous since 2018.
F. Efficiency	Assessment against this criterion has been continuous as a stand-alone criterion since 2014.
G. Monitoring and Reporting	This criterion was changed from 'Monitoring and Evaluation' in 2016 and new sub-categories were defined during 2018-19.
1. Monitoring design and budgeting	The design of monitoring is, since 2018, combined with its budgeting.
2. Monitoring of project implementation	Since 2018, emphasis has been placed on the adequacy and effectiveness of monitoring project implementation rather than the delivery of a monitoring plan. This allows evaluations to give credit to projects that develop a range of ways of monitoring that were not part of the monitoring design.

Evaluation criteria	
3. Project reporting	Assessment against this sub-criterion has been continuous since 2018. It allows evaluations to give full credit to those projects that meet reporting requirements outside of an effective monitoring approach.
H. Sustainability	This criterion focuses on the 'endurance' / 'durability' / longevity of Outcomes.
1. Socio-political sustainability	Assessment against these sub-criteria has been continuous since 2010,
2. Financial sustainability	although the way in which the terms are understood and how likelihood should be assessed has been developed and refined over time.
3. Institutional sustainability	
I. Factors affecting performance and cross-cutting issues	
1. Preparation and readiness	Assessment of this factor has been continuous since 2010 although the category included an assessment of the quality of project design up until 2015.
2. Quality of project management and supervision	Project management and supervision were brought together as one factor from 2018 as the boundaries between the two terms were found to be unhelpful in terms of learning and accountability for non-GEF projects. Requirements for evaluation of GEF projects specify that project supervision (by the Implementing Agency) be assessed separately from project management (by the Executing Agency).
3. Stakeholder's participation and cooperation	Assessment of this factor has been continuous since 2018.
4. Responsiveness to human rights and gender equity	Introduced as a stand-alone factor since 2018. Previously assessed as part of UNEP's policies and strategic priorities.
5. Environmental, social and economic safeguards	Introduced as a stand-alone factor in late 2019. Previously assessed as a negative, unintended effects under Effectiveness.
6. Country ownership and driven-ness	Assessment of this factor has been continuous since 2010.
7. Communication and public awareness	Assessment of this factor has been continuous since 2016.

Across all UNEP project evaluations and reviews, a six-point scale is used to rate performance against each evaluation criterion, as presented in Table 2 below. The rating system and evaluation quality control processes used by UNEP are consistent with those applied by the GEF, United Nations Development Programme (UNDP) and the World Bank. UNEP’s evaluation ratings are regularly benchmarked against those of other relevant organisations.

The UNEP Evaluation Office sets the desired performance level at ‘Satisfactory’. On a six-point scale, ‘Satisfactory’ is the fifth category and, therefore, its lower threshold represents a percentage score of 66.6%. Throughout this report, therefore, commentary is provided on the percentage of project evaluations with performance ratings in the ‘Satisfactory’ or better category. **This means a project has been assessed with either a ‘Satisfactory’ or ‘Highly Satisfactory’ rating.** The abbreviation used for ‘Satisfactory’ or better is S/HS throughout this report.

Table 2. Desired performance level rating scale

Criteria	Rating	Abbrev.	Abbrev.*	Range	Desired Performance
Strategic relevance; Quality of project design; Effectiveness; (Availability of Outputs; Achievement of Project Outcomes; Likelihood of impact); Financial management; Efficiency; Monitoring and reporting; Sustainability; and Factors affecting project performance	Highly Satisfactory	HS	HL	‘Satisfactory’ range	S/HS 
	Satisfactory	S	L		
	Moderately Satisfactory	MS	ML	‘Unsatisfactory’ range	
	Moderately Unsatisfactory	MU	MU		
	Unsatisfactory	U	U		
	Highly Unsatisfactory	HU	HU		

**Sustainability and Impact are rated against a 6-point ‘likelihood’ scale, ranging from ‘Highly Likely’ to ‘Highly Unlikely’.*

To arrive at an overall project performance rating, the UNEP Evaluation Office applies a **weighted scale**⁶ to the individual evaluation criterion ratings. In this weighted scale, 30% of the overall performance score is determined by the rating for ‘Achievement of Project Outcomes’ and a further 20% is determined by the rating for ‘Sustainability’ (i.e. durability of Outcomes). This means that the overall performance rating for all UNEP evaluations is strongly **results-oriented** and places great importance on a project’s effectiveness and the likelihood that the benefits derived from a project will be enduring.

⁶ See Annex 4

Evaluation coverage

The mandate of the UNEP Evaluation Office is to evaluate projects implemented under the Programme of Work. In 2022 and 2023, under the revised evaluation policy, the Evaluation Office sampled a set of projects for independent evaluation to broadly represent the Programme of Work from among the projects reaching operational completion. The projects not selected for evaluation were required to undertake management-led Terminal Reviews following the guidance and templates issued by the Evaluation Office. Completed management-led Terminal Review reports were then subject to a validation and quality assessment process conducted by the Evaluation Office. In this way the performance assessment coverage of UNEP's completing projects is maximised.

In the period spanning from 2010 to 2023, UNEP's Evaluation Office completed over 382⁷ project and programme evaluations. With the current staffing complement, the Evaluation Office has the capacity to complete approximately 30 evaluations each year and this number fluctuates depending on the actual duration of evaluation processes.

Table 3 below shows a breakdown by subprogramme of the number of completed evaluations within the MTS/POW time periods listed.

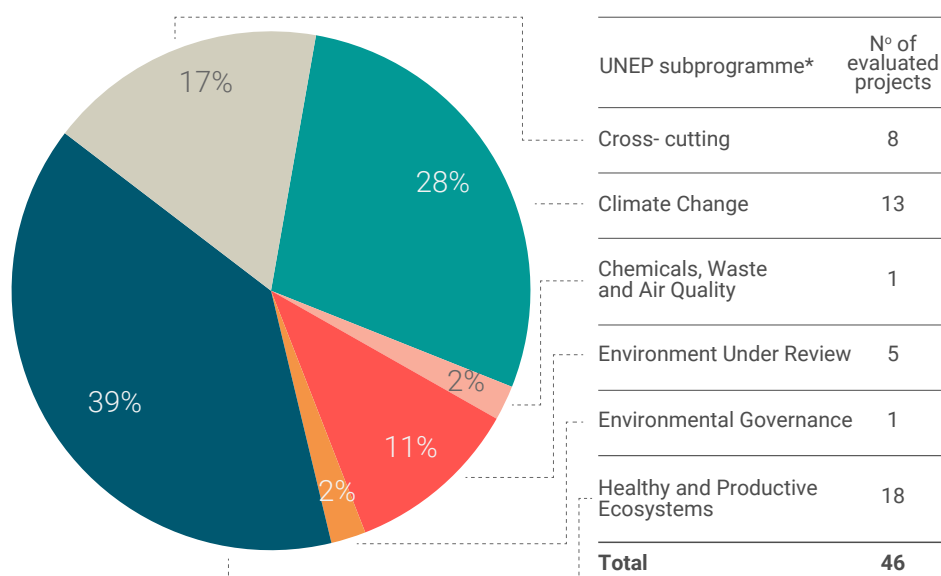
In addition to the individual character of each subprogramme, with some subprogrammes having a small number of projects in their portfolio and others having many, it is noted that the evaluation policy of key donors has for many years played a part in shaping the overall number of project evaluations. Notably, the subprogrammes for Nature Action and Climate Action have received the majority share of GEF grants and the GEF evaluation policy requires a mandatory evaluation. However, with the approval of the new UNEP Evaluation Policy in 2022, the Evaluation Office now selects a sample of projects for independent evaluation to be representative of the Programme of Work. Compliance with the GEF policy for full evaluation coverage of operationally completed GEF projects by UNEP is, nevertheless, achieved because all projects assessed for performance through a management-led Terminal Review process have the final report independently validated by the Evaluation Office prior to submission to the GEF. This practice is in-line with many other GEF agencies e.g. The World Bank, UNDP, IFAD, etc.

⁷ In some cases, evaluations were completed that did not include an evaluation ratings table and are therefore not included in this total.

Table 3. Number of completed evaluations by subprogramme and MTS/POW, 2010-2023

	Climate Change	Resilience to Disasters and Conflicts	Healthy and Productive Ecosystems	Environmental Governance	Chemicals, Waste and Air Quality	Resource Efficiency	Environment Under Review	Cross-cutting	Totals
PoW 2022-2023	13	n.a	18	1	1	n.a	5	8	46
MTS 2018- 2021	28	1	19	19	28	2	2	3	102
MTS 2014-2017	30	4	41	35	7	9	1	2	129
MTS 2010-2013	40	0	40	14	6	5	0	0	105
2010-2023	109	7	116	69	42	20	8	11	382

Figure 1. Distribution of evaluated projects by subprogramme, 2022-2023



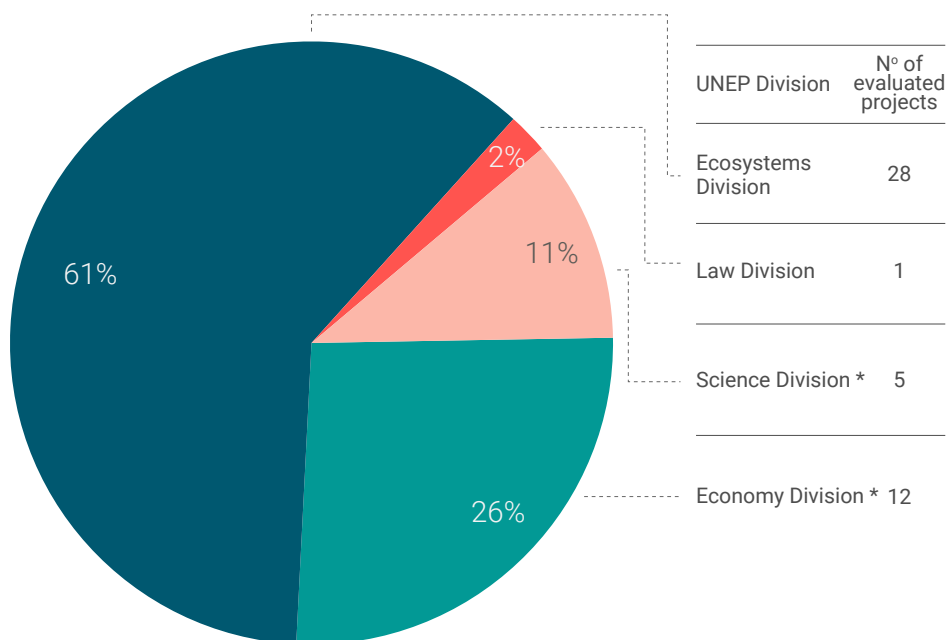
* Name of subprogramme pre-MTS 2022-2025

The Evaluation Office examines its portfolio of project evaluations across several dimensions relevant to the house, namely: funding modality, UNEP Division, global geographic region and Sustainable Development Goals (SDGs). Within each sample there is considerable heterogeneity in the types of projects that have been evaluated (e.g. single country, regional, global, field-based, normative etc.).

In addition, each of the dimensions listed above has some element of complexity: some projects contribute to more than one subprogramme or have been managed by more than one division; the data on funding modality has only been captured on a few named donors; it is assumed that a common definition of a 'global' project is being applied (i.e. any project operating in more than one region or if there is no involvement of specific countries).

The project evaluations completed during 2022 and 2023 with respect to the UNEP division under which they were managed are presented in **Figure 2** below. The number of project evaluations in any given time period is tied to project completions but also reflects the nature and composition of the Divisions themselves; some Divisions have few projects while others have multiple branches/units and manage/oversee many projects within them.

Figure 2. Distribution of evaluated projects by UNEP Division, 2022-2023

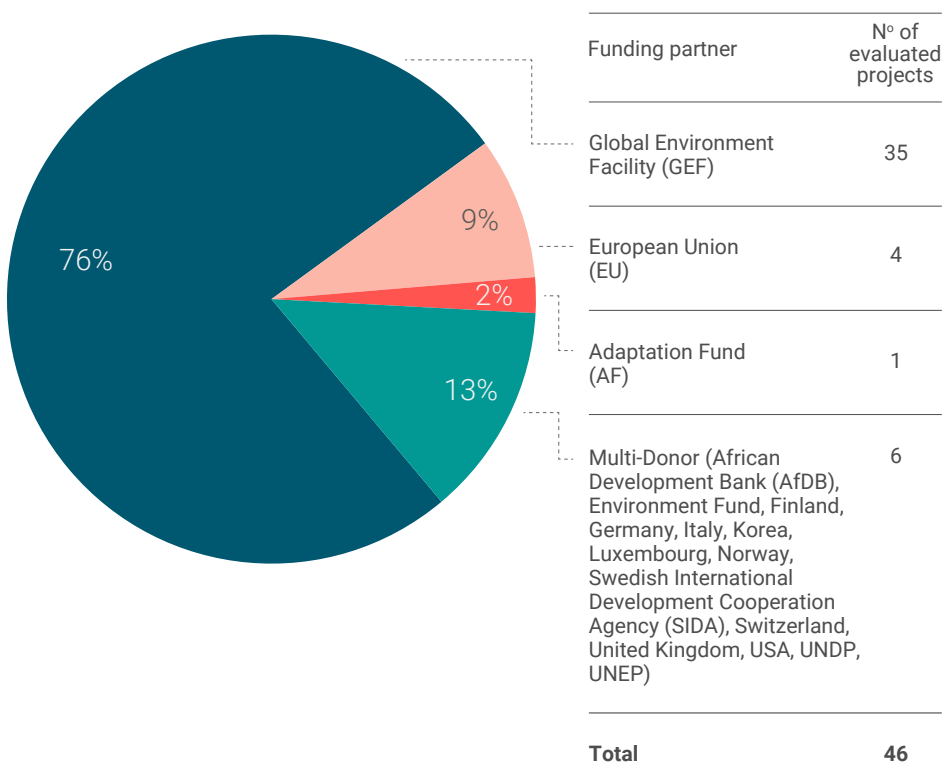


** Economy Division is now named Industry and Economy Division and Science Division is named Early Warning and Assessment Division.*

The number of the project evaluations conducted with respect to the project's funding modality (see **Figure 3**) clearly highlights the effect of the GEF mandatory evaluation requirement as the GEF accounts for 73% of the project evaluations carried out in MTS 2018-2021. Cognisant of this distribution, the Evaluation

Office is making efforts to revise the approach to evaluation in UNEP going forward, such that evaluation coverage becomes more representative of the entire UNEP programme. Details are briefly described in **Chapter 7**.

Figure 3. Distribution of evaluated projects by funding modality, 2022-2023



The proportion of projects evaluated in 2022 and 2023 with respect to the regions where implementation took place is shown in **Figure 4** below. There are overlaps (hence the larger number of instances - 121) as some projects were operating in more than one region.

The Evaluation Office recruits external evaluation consultants to undertake Mid-Term and Terminal Evaluations. The disaggregation by gender of the consultants who completed evaluations in 2020 and 2021, 59% were male and 41% were female (see **Figure 5** below). The UN system-wide Strategy on Gender Parity calls for an inclusive workforce; while the strategy’s goal is 50/50 parity, it is recognized that sustainability at that number is unlikely and therefore parity within the 47%-53% margin may be considered adequate.

Figure 4. Distribution of evaluated projects by geographic region, 2022-2023

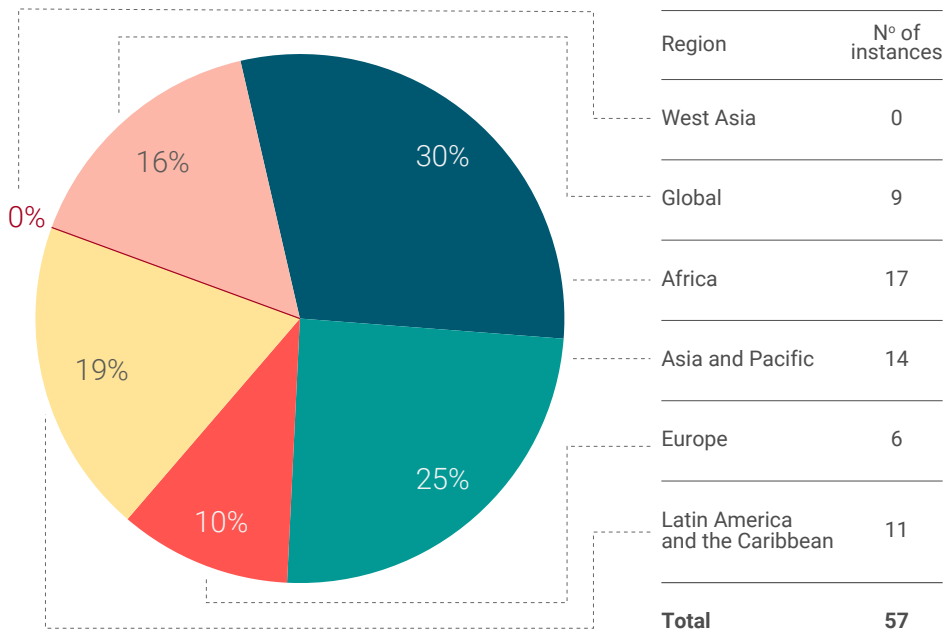
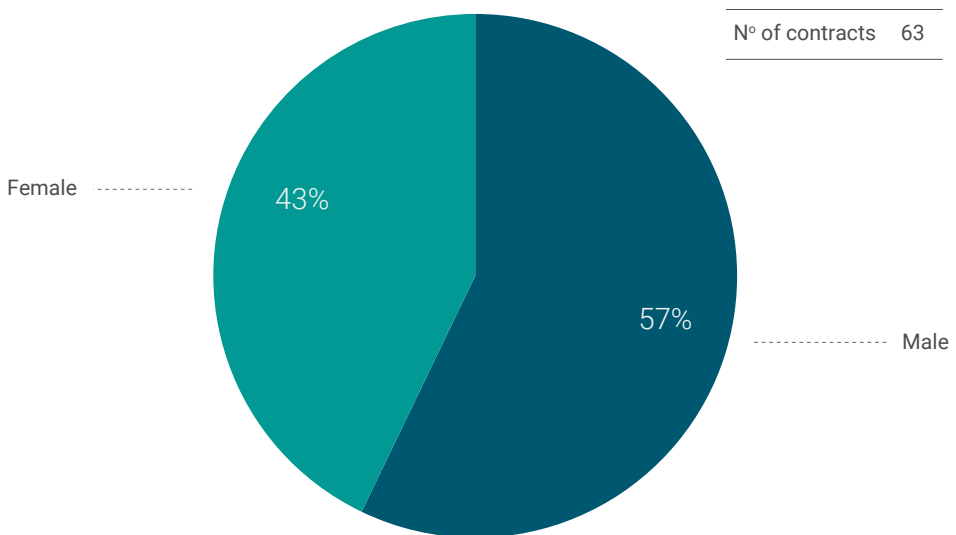


Figure 5. Evaluation consultants by gender, 2022-2023



The strategy also calls for geographic diversity, particularly from under-represented groups, and considers the twin goals of parity and diversity as mutually reinforcing. The consultants referred to in the preceding paragraph were recruited from four geographic sub-regions, as shown in **Figure 6** below.

Regarding the link between a project intervention and the relevant SDGs, it is emphasized that this alignment has been determined retrospectively for some projects and that a significant proportion of projects evaluated in this biennium (47%) were designed and approved before the SDGs came into play. In addition, where a project has been considered as aligned to more than one SDG, all goals have been counted. This underpins the high number of instances of SDGs recorded in Figure 7 below (N=514).

The six SDGs to which UNEP project interventions most frequently refer to were: SDG 1 No Poverty; SDG 7 Affordable and Clean Energy; SDG 12 Responsible Consumption and Production; SDG 13 Climate Action; SDG 15 Life on Land, and SDG 17 'Partnerships for the goals'. There were only two instances of evaluated projects having listed a linkage to SDG 5 on Gender.

Figure 6. Evaluation consultants by geographic representation, 2022-2023

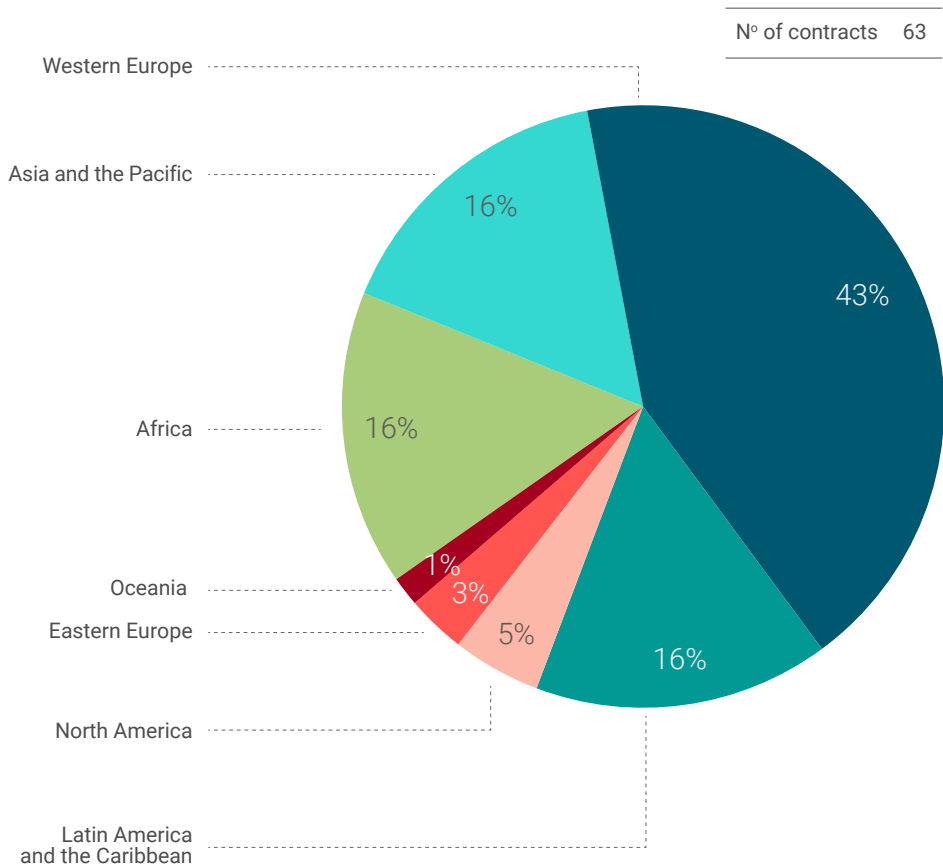
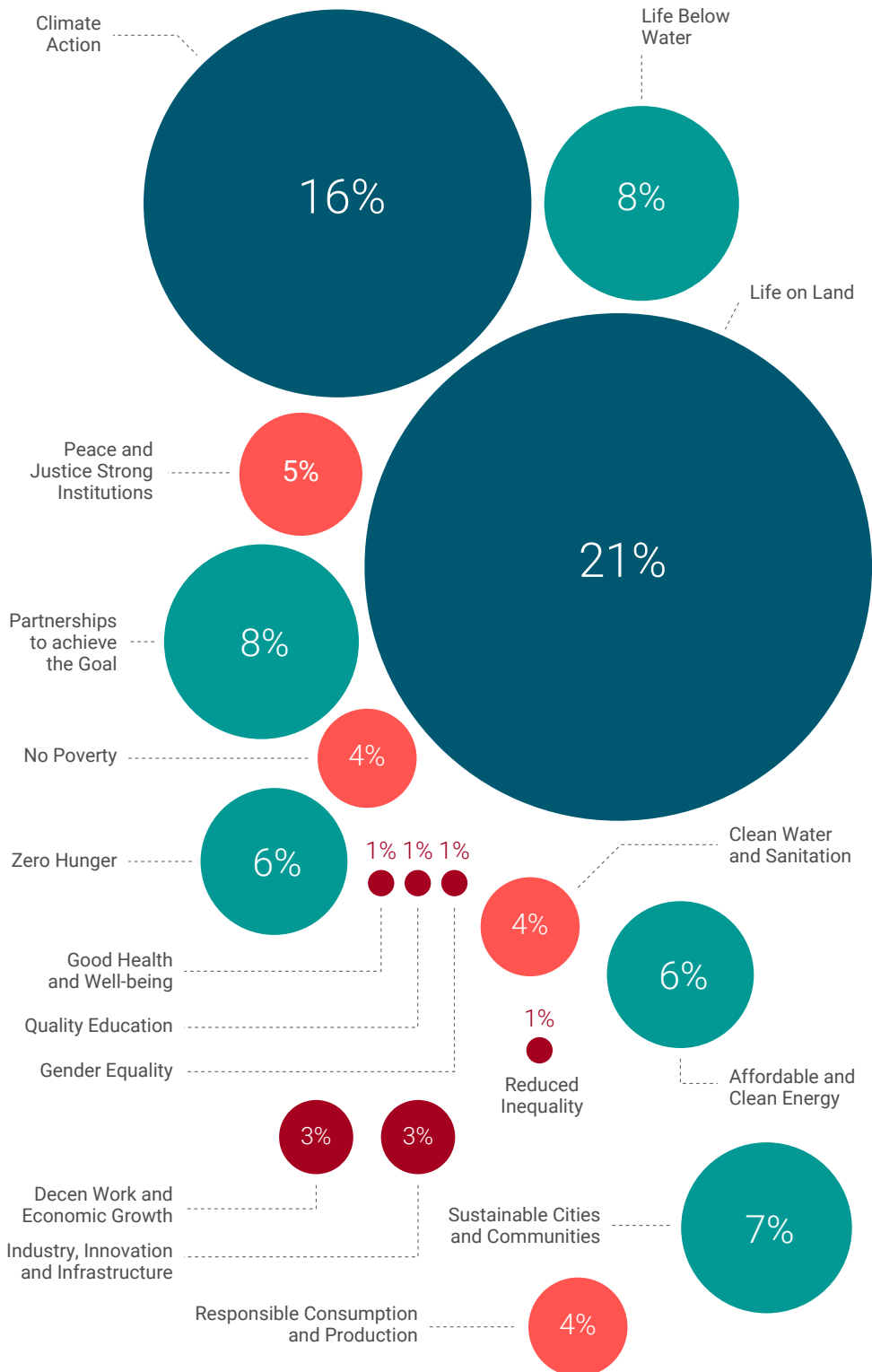


Figure 7. Alignment of evaluated projects to Sustainable Development Goals, 2022-2023



Sustainable Development Goal	No. of instances
SDG 1 No Poverty	5
SDG 2 Zero Hunger	8
SDG 3 Good Health and Well-being	2
SDG 4 Quality Education	2
SDG 5 Gender Equality	2
SDG 6 Clean Water and Sanitation	6
SDG 7 Affordable and Clean Energy	8
SDG 8 Decent Work and Economic Growth	4
SDG 9 Industry, Innovation and Infrastructure	4
SDG 10 Reduced Inequality	1
SDG 11 Sustainable Cities and Communities	5
SDG 12 Responsible Consumption and Production	10
SDG 13 Climate Action	21
SDG 14 Life Below Water	11
SDG 15 Life on Land	28
SDG 16 Peace and Justice Strong Institutions	7
SDG 17 Partnerships to achieve the Goal	11
Total instances	135

Guidance and tools supporting evaluation and management-led Terminal Review processes



Reminder on terminology:
In UNEP, the Evaluation Office has the sole prerogative to issue evaluations. All other assessments overseen by UNEP Managers are considered to be management-led reviews.

In 2022 and 2023, the Evaluation Office prepared and published a new Evaluation Manual⁸. This provides comprehensive guidance on evaluation and review processes. It also provides links, guidance notes, tools and templates to inform evaluation/ review consultants and evaluation / review managers. The aim is to ensure that common approaches to performance assessments are applied across evaluations and management-led reviews to promote transparency, comparability of evaluation/ review ratings and the synthesis of common findings.

⁸ <https://wedocs.unep.org/20.500.11822/42025>

The suite of tools, guidance notes and templates have been extended and updated to support the transition within UNEP to the new project performance assessment model, formalised in the 2022 evaluation policy, where completing projects are assessed through one of two modalities:

- Independent Terminal Evaluations (commissioned and quality assured by the Evaluation Office)
- Management-led Terminal Reviews (commissioned and overseen by Project Managers)

The introduction of management-led Terminal Reviews as a standard form of project performance assessment has created a need for tools, templates and guidance that will ensure review reports meet the requirements of UNEP's different funding partners e.g. Adaptation Fund, GCF, GEF, UNDA. Validation by the Evaluation Office also ensures that project performance ratings are properly benchmarked.

As has been reported in previous biennia, the Evaluation Office liaises with other parts of the house (e.g., the Policy and Programme Division, the GEF Coordination Unit etc.) to ensure that the way in which projects are evaluated is consistent with the guidance UNEP provides during the project design and development process. Such liaison between departments also contributes to an ongoing process of institutional sharing and learning.

With regard to evaluation quality assurance processes, these are applied at inception, draft report and final report stages to ensure: the quality of—and adherence to—the terms of reference, that evaluation processes are in-line with UNEP Norms and Standards and follow UNEP guidance, and that the quality of evaluation reports meets UN system, and internationally agreed, evaluation standards. In addition, UNEP evaluation reports are, periodically, subject to external expert assessments of report quality.

Learning from evaluations

The purpose of evaluations is to meet accountability requirements, inform evidence-based decision-making and contribute to learning. Evidence that learning has been taken up is provided by an assessment of compliance with the recommendations that are derived from evaluations. The Evaluation Office uses a compliance procedure to track the progress of recommendation implementation compliance for a period of 12 months. The Office prepares biannual compliance status reports, and these are reported to senior management at regular intervals. Aggregate compliance data are included in this Evaluation Synthesis Report (see **Chapter 5**). In 2022 and 2023, there has been a trend of increasing compliance

with project-level recommendations, indicative of increased uptake and learning from evaluation findings. In general, however, the evaluation processes themselves are intended to generate learning opportunities.

The Office of Internal Oversight Services (OIOS) holds the mandate for evaluation across the UN Secretariat. In 2022, the Inspection and Evaluation Division of OIOS selected 35 UNEP evaluation reports as examples of good practice to share through its Evaluation Knowledge Management Platform. The Evaluation Office launched a number of knowledge products including a portfolio assessment of lessons learned from thematically related projects on sustainable land management⁹ and collation of commonly issued recommendations issued from evaluations of GEF projects. The Evaluation Office also launched a LinkedIn page that, during the biennium, has grown from zero followers to a total in excess of 9,500, all evaluation reports and knowledge products are shared on this platform. In addition to the lessons learned and recommendations developed during an evaluation, the reports themselves contain considerable insights into different aspects of performance.

Limitations

To fulfil its mandate of evaluating projects forming the Programme of Work, the Evaluation Office is reliant on information regarding the operational completion of projects from all organizational units managing them, and the provision of adequate financial resources from within the projects to support the direct costs of evaluation processes (consultant evaluator fees, travel etc.). During 2022 and 2023, project level information was stored on the UNEP Project Information Management System (PIMS) and, for GEF projects, on the ANUBIS system. In 2023 project level-information of ongoing projects was consolidated into a single system – the Integrated Planning Management and Reporting System (IPMR). In future biennia, it should provide comprehensive and reliable information on all projects under implementation and nearing completion, greatly aiding evaluation planning.

Given the manual nature by which the list of projects eligible for evaluation have, to date, been compiled, there is a possibility of a positive bias in the projects the Evaluation Office evaluates. Completing a project on time and ensuring that an adequate evaluation budget is available by the time it is needed are both characteristics of good project management. It is logical to assume that strong project management skills are likely to be associated with stronger project performance. By contrast projects that do not complete on time, which do not

⁹ <https://wedocs.unep.org/20.500.11822/43560>

budget for evaluation, or whose managers fail to inform the Evaluation Office of their completion, will tend to be under-represented in the sample of projects evaluated.

While the overall sample size (46 evaluations in the current biennium) forms a robust basis for learning lessons, when this sample is disaggregated by dimensions of interest to the house (e.g. subprogrammes, their financial value, or projects implemented in Africa, etc.), its reliability for providing insights is more limited because of the small number of projects involved in any sub-category.

Chapter 3

Analysis of Project Performance in 2022 and 2023

In 2022 and 2023, a total of 48 evaluations were completed for projects with start dates ranging between 2009 and 2019, the majority of these projects (82%) started between 2013 and 2017. The data used for the analysis in this Chapter cover the 46 projects for which performance ratings by criteria were included in the evaluation reports¹⁰. Performance trends across various evaluation criteria are presented and analysed based on data from evaluation reports completed over a span of at least five biennia. In some instances, the narrative includes shared learning derived from recommendations and best practice examples from past evaluations of UNEP projects.

¹⁰ In some instances, programme evaluations and special studies do not confer any criteria-based performance ratings, so references to the number of completed evaluations may vary within this report.

Since project interventions are very diverse thematically, geographically, and in terms of their resource envelopes, conscious effort is needed to ensure that performance assessments are made in a consistent manner. In this regard, the Evaluation Office is continually aiming to improve the objectivity and comparability of its evaluation approach across UNEP projects. External evaluation consultants are routinely provided with standard Terms of Reference for project evaluations, as well as detailed guidance on evaluation processes and methods, report structure, content, and quality, to help maintain consistent approaches and performance assessments.

To reduce the variability that could be introduced across the work of a large number of individual consultants and evaluation processes, and to improve the transparency of evaluative judgements, the Evaluation Office provides a Criterion Ratings Description Matrix as a basis for making the determination of evaluation ratings more explicit. This matrix presents the evidence requirements to support each performance level under each criterion.

In the figure below (**Figure 8**) a summary of project performance by criteria is presented based on data collated from the evaluation reports completed in the 2022-23 biennium and using a six-point ratings scale from Highly Satisfactory down to Highly Unsatisfactory.

UNEP Evaluation Office sets the minimum **desired performance level at 'Satisfactory'** and focuses its attention on the percentage of cases where ratings have been awarded at the levels of 'Satisfactory' or better (i.e., either 'Satisfactory' or 'Highly Satisfactory', abbreviated to S/HS).

A brief overview of the data in **Figure 8** above shows that the evaluation criteria with the highest proportion of projects (at least three quarters) achieving 'Satisfactory' or better (S/HS) ratings include: Strategic Relevance (96%); Availability of Outputs (83%); Financial Management (82%); and Stakeholder Participation and Cooperation (78%).

The assessment of some evaluation criteria however, has revealed sub-optimal performance where at least 15% of projects evaluated against those criteria achieved ratings in the 'Unsatisfactory' range (i.e., 'Moderately Unsatisfactory', 'Unsatisfactory' or 'Highly Unsatisfactory'). These criteria include: Monitoring and Reporting (30%); Sustainability (27%); Responsiveness to Gender and Human Rights (20%); Preparation and Readiness (17%); and Quality of Project Management and Supervision (15%).

Figure 8. Summary of project performance against evaluation criteria, 2022-2023

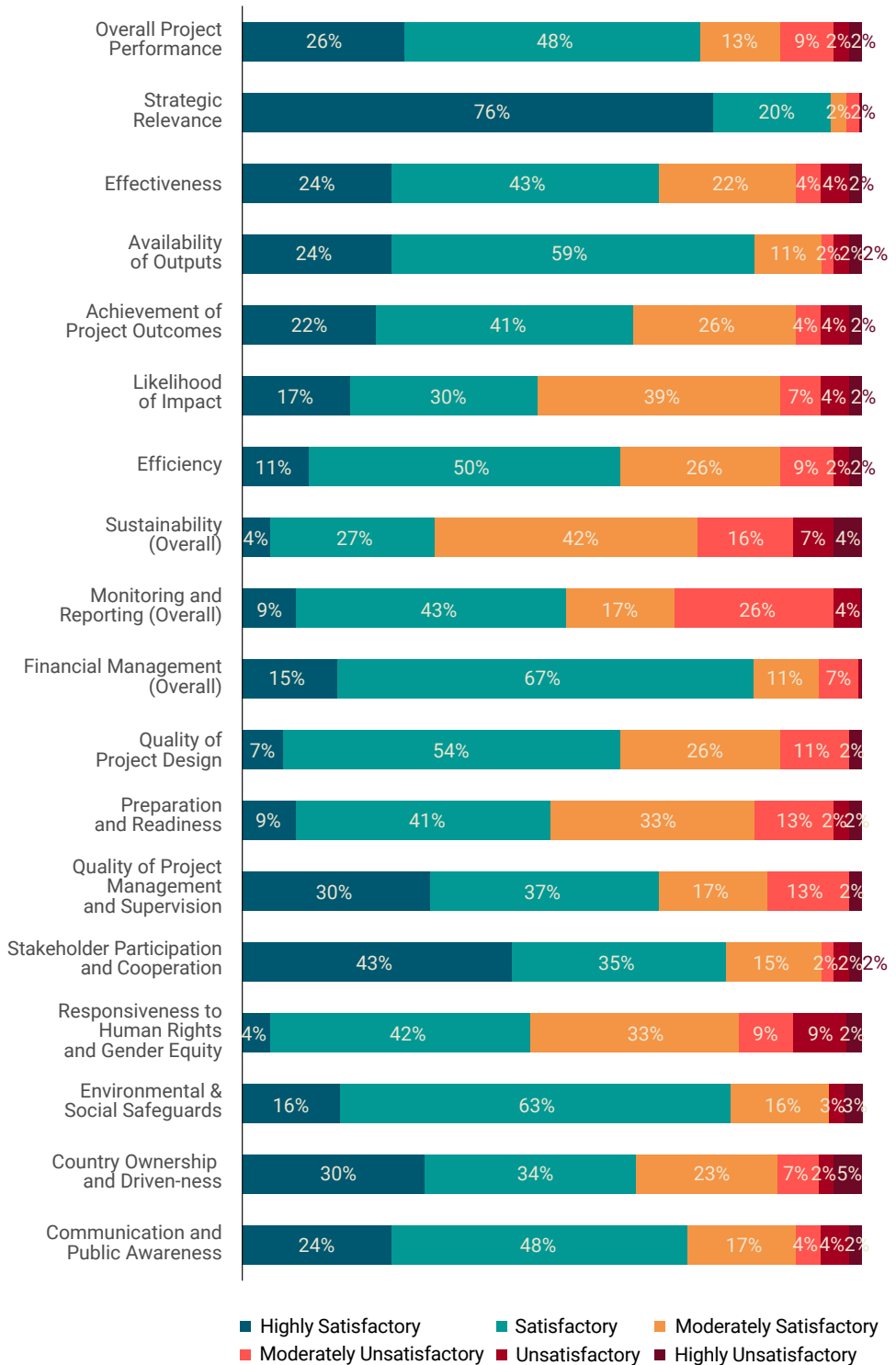


Table 4 below shows a comparison in project performance across the main evaluation criteria, between projects evaluated in 2022-2023 and the previous biennium.

Table 4. Comparison of projects attaining S/HS ratings in the main evaluation criteria in 2020-2021 and 2022-2023

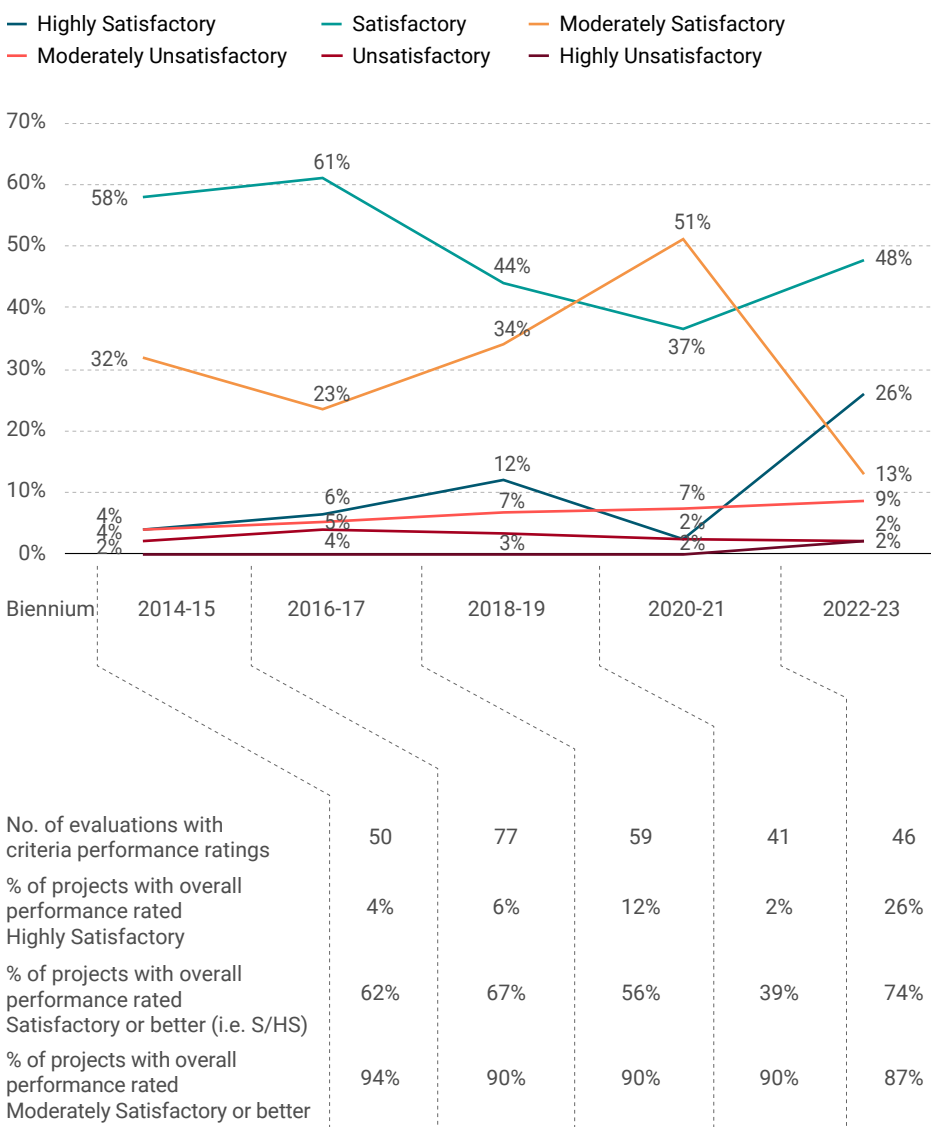
Evaluation criteria	2020-2021 Percent of evaluated projects rated S/HS (n=42)	2022-2023 Percent of evaluated projects rated S/HS (n=46)
Overall Project Performance	39%	74%
Strategic Relevance	96%	96%
Quality of Project Design	52%	61%
Effectiveness	39%	67%
Availability of outputs	71%	83%
Achievement of outcomes	34%	63%
Likelihood of impact	34%	48%
Financial Management	46%	83%
Efficiency	51%	61%
Monitoring and Reporting	49%	52%
Sustainability	10%	31%

The data show that higher proportions of evaluated projects achieved S/HS ratings across all criteria in the sample for 2022 and 2023 compared with those evaluated in the previous biennium. A noteworthy increase in proportion of projects attaining S/HS ratings across the 2022-23 sample is for the criterion on Financial Management; an increase of 37 percentage points when compared to the sample for the previous biennium (46% in 2020-21; 83% in 2022-23). Likewise, there have been other increases in the 2022-23 sample for the following criteria: Achievement of Outcomes (29 percentage point increase); Effectiveness (28 percentage point increase); and Sustainability (21 percentage point increase).

Overall project performance

Figure 9 below shows the general trend in overall project performance over a 10-year period since the 2014-15 biennium. The 2022-23 biennium shows the highest proportion of projects rated Highly Satisfactory for their overall performance, with 26% of the rated projects reaching this level. In 2022 and 2023, 74% projects attained a rating of S/HS for their overall performance, as compared to an average of 56% across the four previous biennia (62% in 2014-15; 67% in 2016-17; 56% in 2018-19; and 39% in 2020-21).

Figure 9. Comparison of overall project performance, 2014-2023

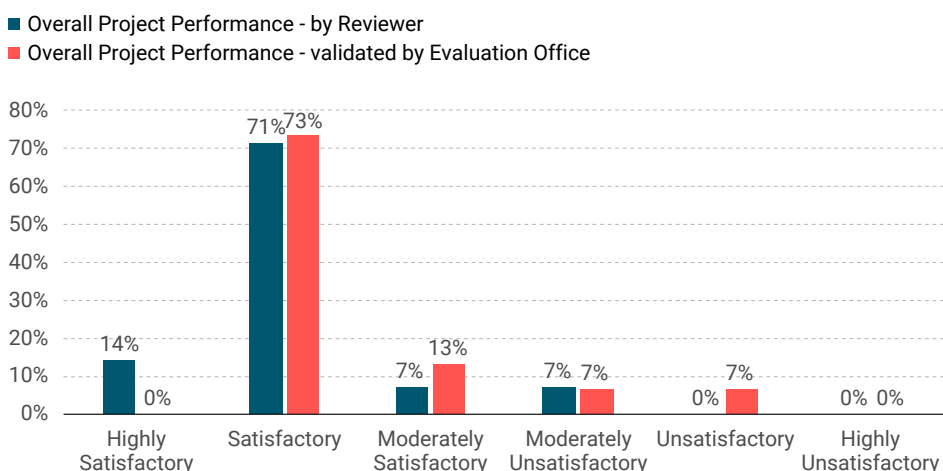


Overall project performance in management-led Terminal Reviews

In 2022, UNEP approved a new evaluation policy which introduced a new performance assessment modality, namely management-led Terminal Reviews. In 2022 and 2023, a small group of 15 management-led Terminal Review reports were completed and submitted to the Evaluation Office for independent validation and an assessment of the report's quality. These projects were rated against the same performance criteria as used in independent terminal evaluations.

The chart below (Figure 10) shows the distribution of ratings for the criterion Overall Project Performance among the projects whose performance were assessed through management-led processes. The figure includes the comparison between the Reviewer's assessment of the performance rating and the validated rating approved by the Evaluation Office. The data show that the majority (71% and 73% respectively) of these projects were rated at the 'Satisfactory' level for overall performance by both the Reviewer and the Evaluation Office.

Figure 10. Overall project performance in management-led Terminal Reviews, 2022-2023



Overall project performance presented by project size

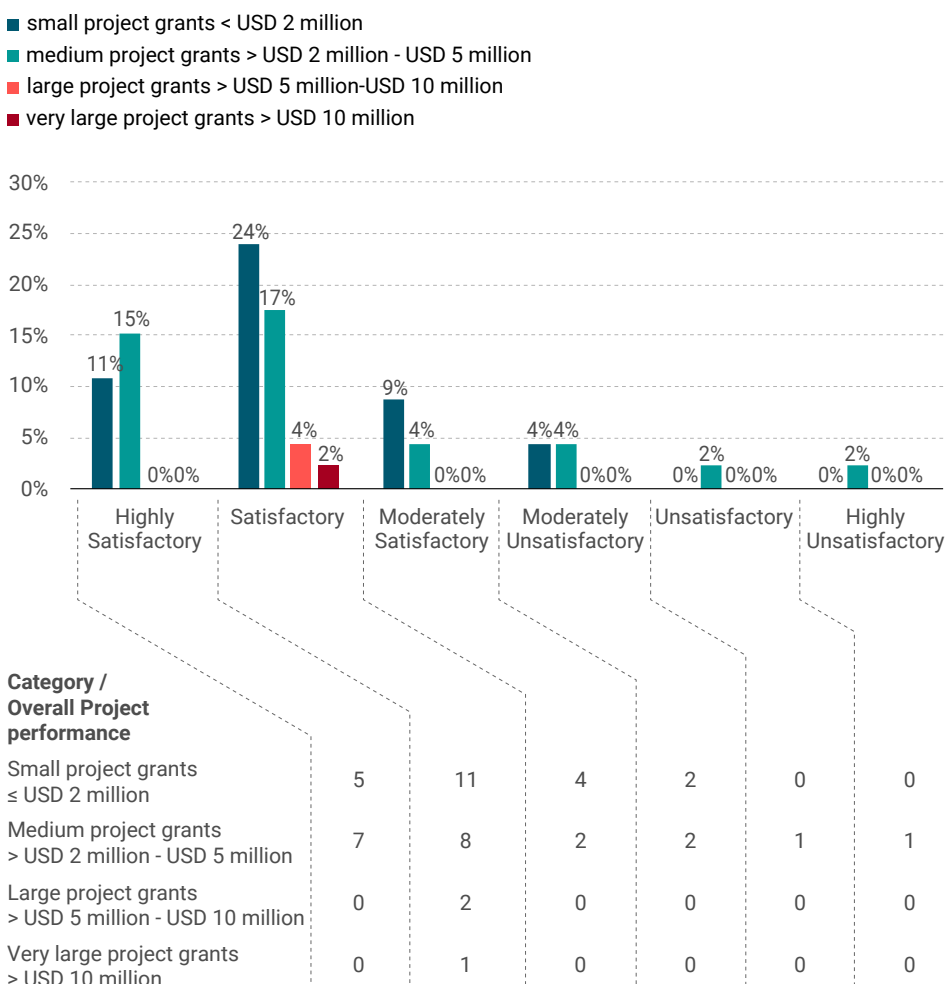
The sum total (in USD) of project grants recorded in evaluation reports completed in 2022 and 2023 is > USD 157 million¹¹. Figure 11 below shows the overall project performance, categorised according to the secured funding indicated in the projects' approved budgets, as follows: (i) small project grants

¹¹ Co-finance amounts (in-cash and in-kind) have not been included due to inconsistencies that make it difficult to verify the actual co-funded amounts.

of < USD 2 million; (ii) **medium** project grants of > USD 2 million and < USD 5 million; (iii) **large** project grants of > USD 5 million and < USD 10 million; and (iv) **very large** project grants of > USD 10 million.

Figure 11 below, has insufficient numbers of projects in each financial size category to draw any meaningful conclusions. However, the differential performance of UNEP projects in relation to their financial envelope will be tracked over future biennia.

Figure 11. Overall project performance by project size, 2022-2023

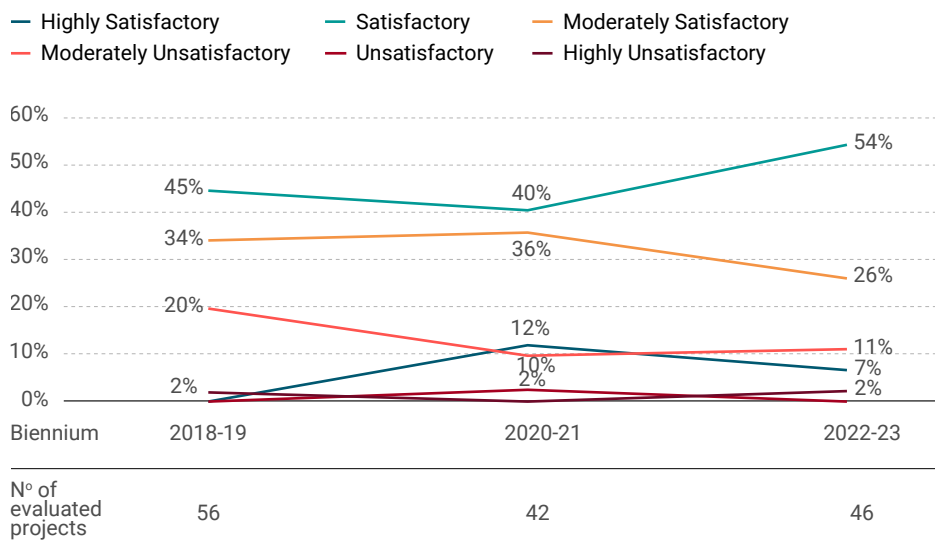


Quality of project design

Although this criterion has only been rated since 2017, the Evaluation Office has been assessing and documenting project design quality since 2011. The quality of project design is assessed during the evaluation inception phase against

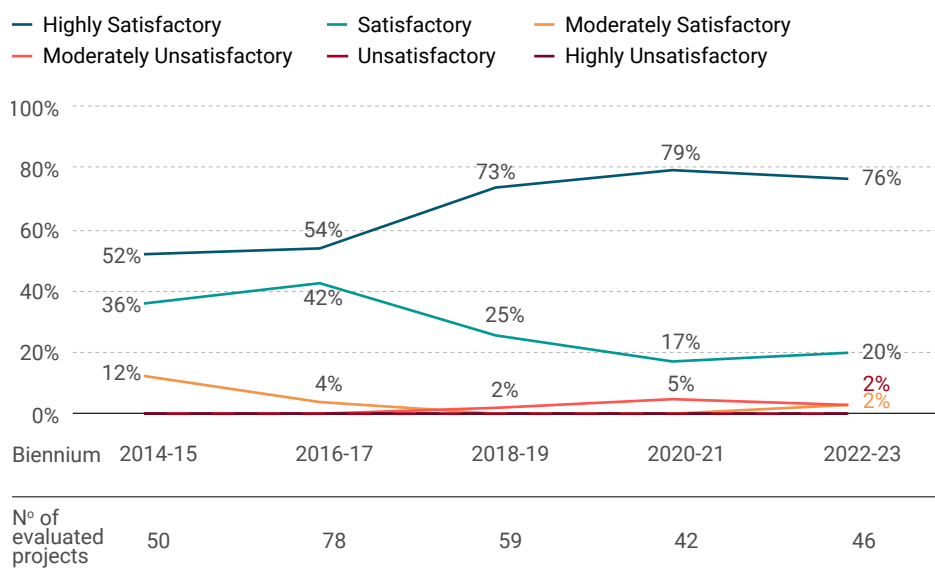
various criteria, and an overall Project Design Quality rating is established using a weighted scoring approach. In **Figure 12** below it shows an increase in the proportion of project designs achieving a S/HS rating over the three (biennia, from 2018-19 (45%) to 2022-23 (61%). A more detailed analysis is presented in **Chapter 6**.

Figure 12. Quality of project design, 2018-2023



Strategic relevance

Figure 13. Comparison of strategic relevance (overall), 2014-2023



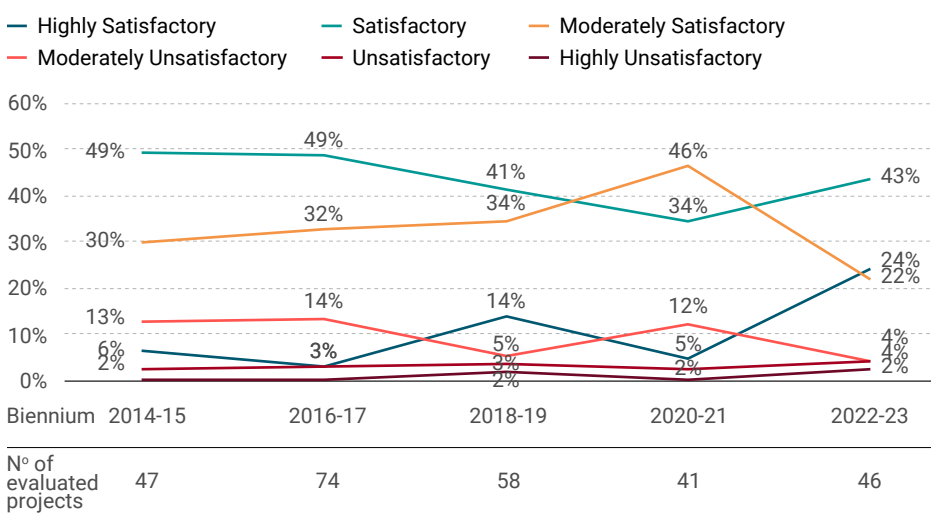
This criterion assesses alignment with strategic frameworks and priorities at the levels of UNEP, donors, regions and countries, as well as including a sub-category on complementarity, which corresponds with the OECD-DAC criterion Coherence. Complementarity refers to how well the project, either at design stage or during the project inception or mobilization, took account of ongoing and planned initiatives that address similar needs of the same target groups. Strategic relevance remains the evaluation criterion with a consistently high proportion of S/HS ratings in every biennium. As shown in **Figure 13** above, up to 96% of the projects evaluated in the 2022-23 biennium achieved a S/HS rating, as compared to an average of 94% across the four previous biennia. In 2022 and 2023, 76% of projects were rated as 'Highly Satisfactory' against this criterion.

Effectiveness

The assessment of Effectiveness consists of three sub-criteria (Availability of Outputs; Achievement of Outcomes and Likelihood of Impact), and accounts for 45% of the overall project performance weighted score (5%, 30% and 10% respectively). Performance ratings for the Effectiveness criterion therefore have a large influence on the overall project performance rating.

As shown in **Figure 14** below, projects evaluated in 2022 and 2023 showed an increase in the proportion rated S/HS for their overall effectiveness, with 67% of the sample as compared to an average of 50% across the four previous biennia (55% in 2014-15; 52% in 2016-17; 55% in 2018-19; and 39% in 2020-21). The highest proportion of projects (24%) achieving a 'Highly Satisfactory' performance rating for criterion was also recorded in the 2022-2023 biennium.

Figure 14. Effectiveness (overall), 2014-2023



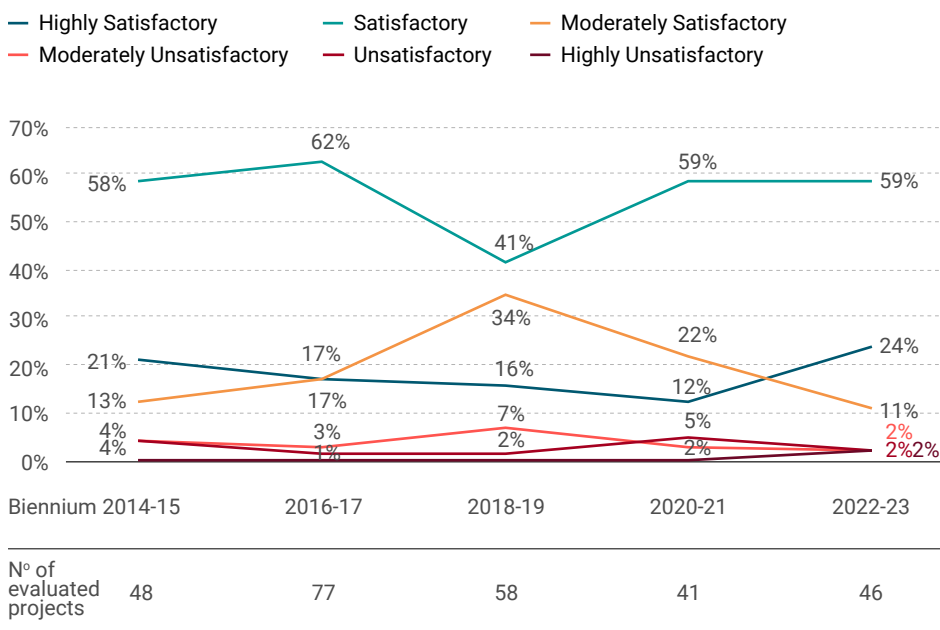
The three sub-criteria under the Effectiveness criterion are discussed in sections 3.5 - 3.7 below.

Availability of outputs

UNEP defines an output as the “availability (for intended beneficiaries/users) of new products and services and/or gains in knowledge, abilities and awareness of individuals or within institutions”. Since 2018, project evaluations have more clearly and consistently considered the following elements in the assessment: proportion of planned outputs made available to the target users/beneficiaries; timeliness and suitability for purpose of those outputs most important for the achievement of outcomes; and quality and utility of the outputs to identified users. The Evaluation Criteria Ratings Matrix¹², is a tool used in all project evaluation, it describes these factors and aids consistent assessments across project evaluations.

There is a relatively high proportion of projects attaining S/HS ratings for this criterion (83%) in the 2022-23 biennium, as compared to an average of 71% across the last four biennia (see Figure 15).

Figure 15. Comparison of availability of outputs by biennium, 2014-2023



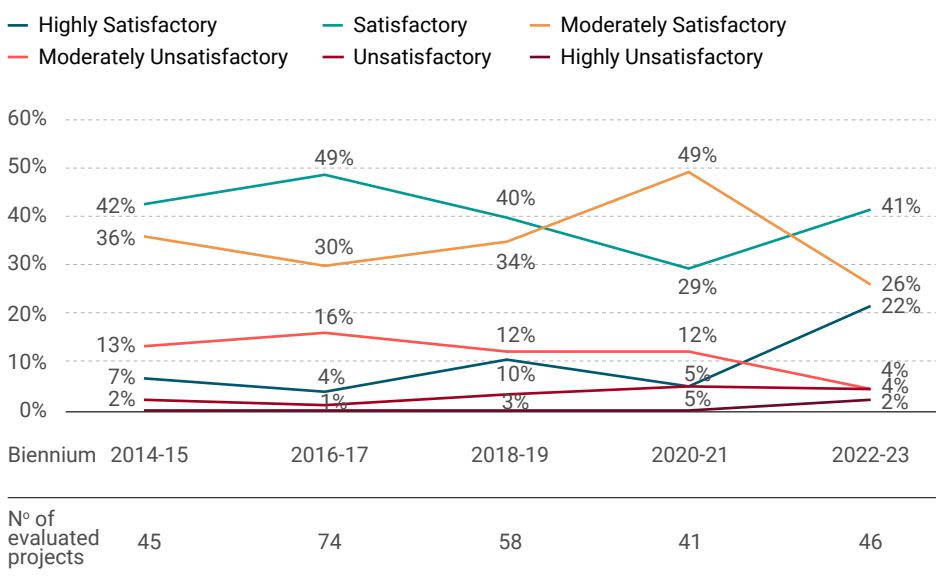
¹² Tool #02 <https://communities.unep.org/display/EQU/INDEPENDENT+EVALUATION+TOOLS+AND+TEMPLATES>

Achievement of project outcomes

According to UNEP¹³, outcomes refer to the use (i.e. uptake, adoption, application) of outputs by intended beneficiaries, observed as changes in their behaviours, attitudes or conditions¹⁴. Evaluations assess the level of achievement of expected outcomes, and whether those achievements can be demonstrated. Outcomes are also the points in the change process at which agency¹⁵ is clearly transferred from those who are responsible for delivering the project, to those who are expected to benefit from, or use and take forward, the projects' results.

Project performance against this criterion in 2022 and 2023, shows an increased proportion of the projects achieving a S/HS rating (63%) as compared to the average across the four previous biennia of 47% (49% in 2014-15; 53% in 2016-17; 50% in 2018-19; and 34% in 2020-21), (Figure 16). The 2022-23 biennium also records the highest proportion of projects (22%) achieving a 'Highly Satisfactory' rating in the 10-year period (2014-2023).

Figure 16. Achievement of outcomes, 2014-2023



A comparison between Figure 15 and Figure 16 above, shows in the sample for 2022 and 2023 of evaluated projects that the proportion of S/HS ratings for Achievement of Outcomes (63%) is distinctly lower than the ratings for the Availability of Outputs (83%). Project performance at outcome level is a

¹³ UNEP Glossary of results definitions

¹⁴ UNEP formalized its own articulation of the definition of a project outcome, along with other results terminology, in 2019.

¹⁵ Refers to the capacity, condition, or state of acting or of exerting power (Merriam-Webster definition)

clear indication of the **effect** of UNEPs work, and the difference in the S/HS ratings between the Availability of Outputs and the Achievement of Outcomes, as shown by these data, is an area that continues to require management attention¹⁶ both in terms of the attention paid to the design phase (including adequate budgets/resource allocation to support processes beyond output level) and during implementation (monitoring of assumptions and drivers and outcome-oriented adaptive management).

Table 5. Comparison of performance between availability of outputs and achievement of project outcomes, 2014-2023

	S/HS 2014-2015	S/HS 2016-2017	S/HS 2018-2019	S/HS 2020-2021	S/HS 2022-2023
Availability of Outputs	79%	79%	57%	71%	83%
Achievement of Project Outcomes	49%	53%	50%	34%	63%
Percentage point difference	30 points	26 points	7 points	37 points	20 points

Likelihood of impact

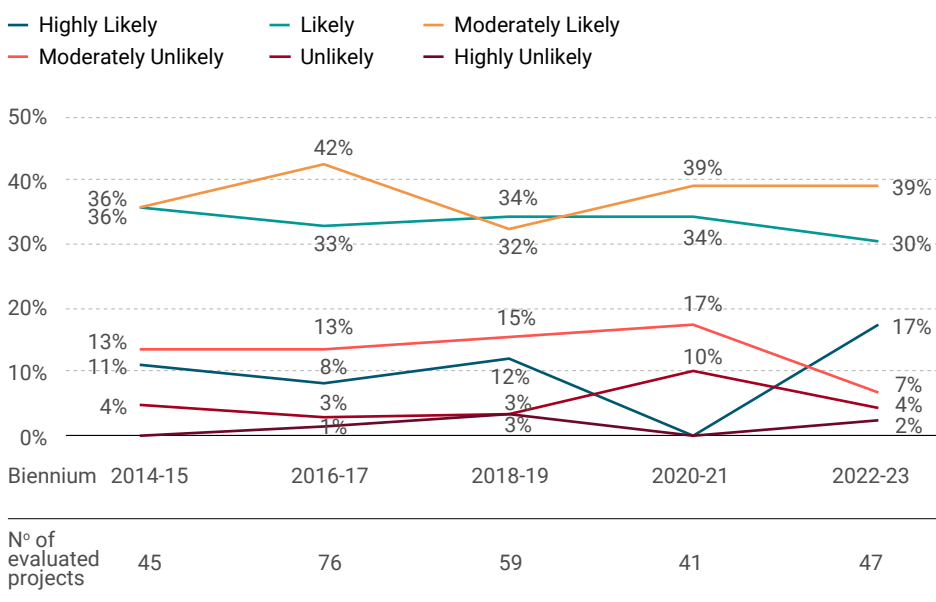
This criterion considers the extent to which outcomes have been achieved, and assumptions¹⁷ and drivers¹⁸ are being seen to hold (i.e. are contributing to the projects' results emerging or developing). It is recognized that Impacts, by definition long-lasting positive effects (and therefore stable changes of state), are not under the sole influence of a project. Further, evaluations are undertaken as soon after a project's operational completion as is feasible, thus it is often too early to gather clear evidence of impacts. Evaluations, therefore, assess the *likelihood* that the intervention may lead, or contribute to, the intended impacts (and in some cases, unintended negative effects). The Evaluation Office supports this assessment process by providing structured guidelines and tools.

¹⁶ This issue has been raised in previous biennial evaluation synthesis reports (2018-2019 and 2020-2021)

¹⁷ An assumption is a significant external factor or condition that needs to be present for the realization of the intended results but is beyond the influence of the project and its partners.

¹⁸ A driver is a significant external factor that, if present, is expected to contribute to the realization of the intended results of a project. Drivers can be influenced by the project and its partners.

Figure 17. Likelihood of impact, 2014-2023



The proportion of projects achieving L/HL¹⁹ ratings against this criterion was greater (47%) in 2022 and 2023 than in the three previous biennia as shown in **Figure 17** above. It was also greater than the average of 42% across the four previous biennia (47% in 2014-15, 41% in 2016-17; 46% in 2018-19; and 34% in 2020-21). In 2022 and 2023, there was a smaller proportion of projects that were assessed to be within the ‘Unlikely’ range to achieve their intended impacts (13% ‘Moderately Unlikely’, ‘Unlikely’ or ‘Highly Unlikely’), as compared to previous biennia (17% in both 2014-15 and 2016-17; 21% in 2018-19 and 27% in 2020-21).

Sustainability

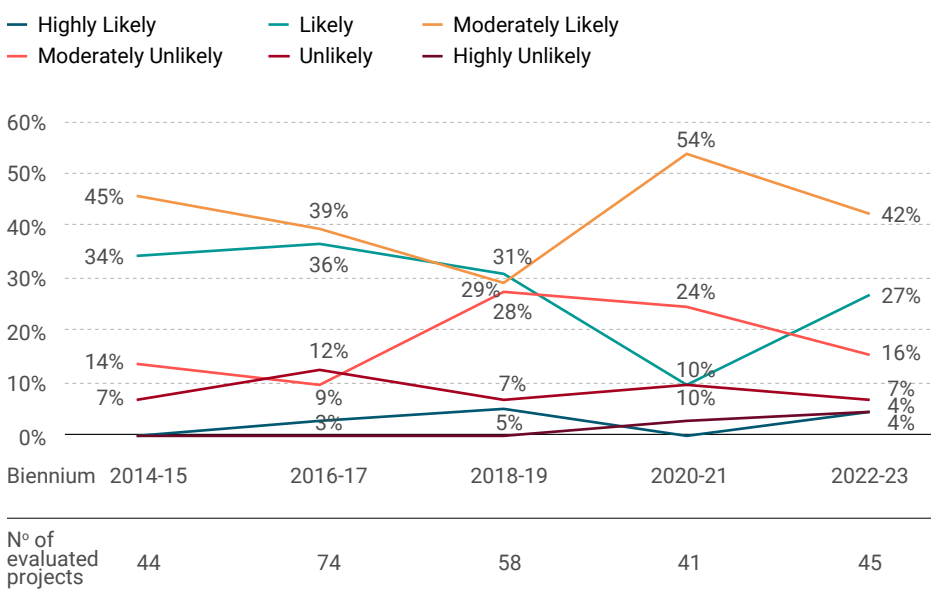
The assessment of sustainability considers the key conditions or factors that are likely to either undermine or support the persistence of project benefits after the external funding and assistance ends. It aims to determine if an appropriate exit strategy has been put in place, to ascertain to what extent follow-up work is necessary, has been planned or initiated, and includes an assessment of any measures taken to mitigate risks that might threaten the endurance of project benefits.

¹⁹ Likelihood of Impact and Sustainability are both rated against a six-point rating scale, labelled as: Highly Unlikely; Unlikely; Moderately Unlikely; Moderately Likely; Likely; Highly Likely. Abbreviations are therefore: HU, U, MU, ML, L and HL. Likely (L) and Highly Likely (HL) abbreviated as L/HL equate to Satisfactory (S) and Highly Satisfactory (HS) abbreviated as S/HS.

Projects often lack mechanisms or budget for monitoring results beyond the funded implementation period. Where no clear exit strategy has been developed, there will also be no clear process for the replication and/or upscaling of the intervention. Even in cases where a project has demonstrated to a certain degree that higher-level / longer-term outcomes have been/are likely to be achieved, the momentum can only be sustained if there is an exit strategy that also allows for replication/scaling up to other parts of the country/region or in similar contexts.

In 2022 and 2023, 31% of evaluated projects were rated as ‘Likely’ or ‘Highly Likely’ to sustain project benefits as compared to the average of 30% over the previous four biennia (34% in 2014-15, 39% in 2016-17, and 36% in 2018-19). As shown in Figure 18 below, there was a reduction in the proportion of projects rated ‘Likely’ or ‘Highly Likely’ to sustain project results over time in 2020 and 2021, with only 10% reaching this performance level.

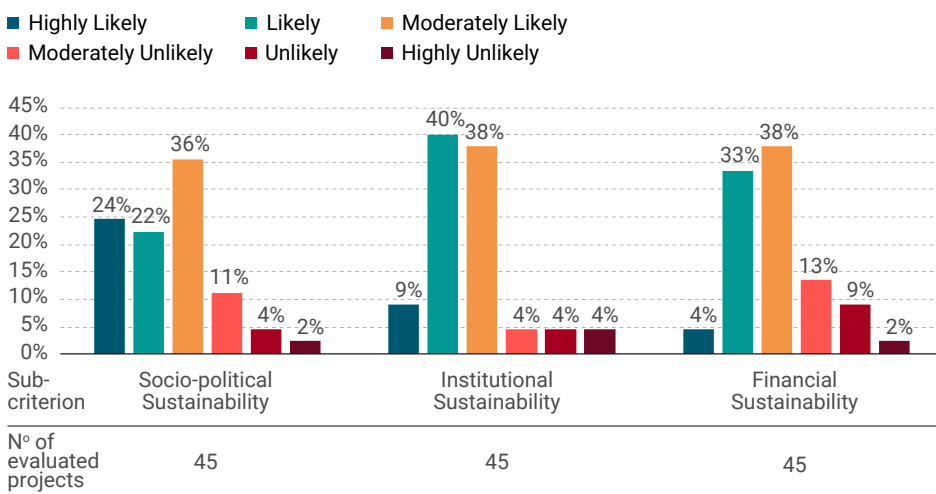
Figure 18. Sustainability (overall), 2014-2023



In calculating overall project performance, the Evaluation Office applies a weight of 20% to the Sustainability criterion, which is allocated across the following sub-criteria: **socio-political, institutional and financial sustainability**. Conceptually, these three sub-criteria are regarded as interrelated limiting factors, therefore the overall rating for sustainability is always set at the lowest rating achieved across these sub-criteria. Project performance in 2022 and 2023 is presented in Figure 19 below.

At the level of the sub-criteria, evaluated projects have shown stronger performance in socio-political sustainability and institutional sustainability (with L/HL ratings in 46% and 49% of projects respectively), and weaker performance in financial sustainability (L/HL ratings in 37% of projects).

Figure 19. Project performance in socio-political, institutional, and financial sustainability, 2022-2023



Socio-political sustainability

Under the first sub-criterion, evaluations assess the extent to which social or political factors are required to support the continuation of project outcomes, including the level of ownership, interest and commitment among government and other stakeholders to take the project achievements forwards. Generally, socio-political sustainability is enhanced when projects are deemed to be relevant, stakeholders (including those at the community level) actively participate, and the project works closely with national and local governments to create a conducive policy environment. Among the factors that have been found to negatively affect socio-political sustainability include political unrest, shifting political and economic priorities, and lack of local political goodwill. In 2022 and 2023, the proportion of evaluated projects rated ‘Likely’ or ‘Highly Likely’ for this sub-criterion was 46% (Figure 19).

Institutional sustainability

This sub-criterion of Sustainability gauges the likelihood that project outcomes will be sustained by the existing institutional and legal frameworks, policies, governing structures, and processes. This largely depends on the degree to which key stakeholders are involved in project execution, the extent to which the project contributes to defining clear governance and institutional frameworks that can remain in place after projects completion, as well as how robust the

relevant institutional features are. In 2022 and 2023, 49% of evaluated projects were rated 'Likely' or 'Highly Likely' in this sub-criterion (Figure 19).

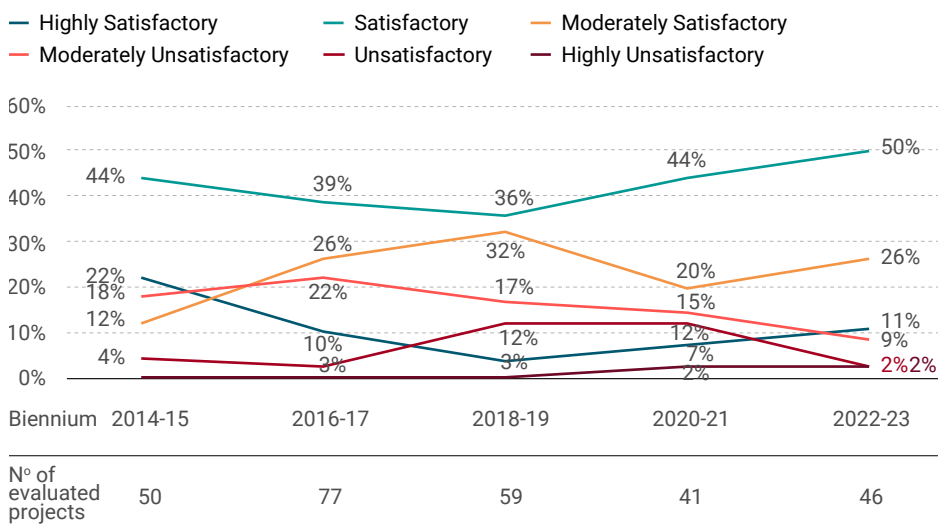
Financial sustainability

Evaluations assess the extent to which the continuation of project results, and the eventual impact of the project are dependent on financial resources²⁰. Evaluations also assess the likelihood that adequate financial resources will be available to use the capacities built by the project, and whether there are any financial risks that may jeopardize continuance of project results and onward progress towards Impact. As evidenced by the data, this is the dimension of the Sustainability where constraints are most commonly experienced. Thirty-seven percent (37%) of the projects evaluated against this sub-criterion in the current biennium were assessed as being 'Likely' or 'Highly Likely' to financially sustain results beyond the funded implementation period (Figure 19).

Efficiency

Under this criterion, the evaluation undertakes an assessment of the **cost-effectiveness** (the extent to which an intervention has achieved, or is expected to achieve, its results at the lowest possible cost) and the **timeliness** (whether planned activities were delivered according to expected timeframes as well as whether events were sequenced efficiently) of project execution.

Figure 20. Efficiency, 2014-2023



²⁰ Such resources can be from multiple sources, such as the national budget, public and private sectors, development assistance, etc.

So-called 'no-cost' extensions are considered under Efficiency because they reflect an increase in project support costs when the implementing timeframe is extended. Evaluations assess to what extent such extensions were justified, or could have been avoided, and identify any negative impacts they might have caused. Qualitative observations from recent evaluation experiences suggest that matching the project timeframe and budget (at design) more realistically to the nature of activities to be implemented, and more efficient (i.e. reduction of delays) in project start-up phases could help to avoid many project extensions.

As illustrated above (**Figure 20**) the proportion of projects at the S/HS performance level was recorded at 61% in 2022-23 biennium, as compared to the average of 51% over last four biennia (66% in 2014-15, 49% in 2016-17, 39% in 2018-19, and 51% in 2020-21). This biennium also had the lowest proportion of projects with ratings in the 'Unsatisfactory' range (MU-HU) with 13% of the evaluated projects falling within that performance category in the 10-year period from 2014 to 2023.

Financial management

In 2016, financial management was established as a stand-alone evaluation criterion. It is assessed under three sub-criteria: **adherence** to UNEP's financial policies and procedures (whether proper financial management standards and UNEP's financial management policies were adhered to); **completeness** of financial information (whether standard financial documentation is available, complete or accurate); and **communication** between finance and project management staff (whether the level of communication between the Project Manager and the Fund Management Officer has supported the effective delivery of the project).

The proportion of S/HS ratings under this criterion was 82% in 2022 and 2023, with an increase of 36 percentage points as compared to the previous biennium, in which only 46% of projects were assessed to have performed to that level (see **Figure 21** below). This is the highest proportion of projects achieving a rating of S/HS across all the biennia since 2014 and exceed the average of 56% over the last four biennia.

Below (**Figure 22**) is a breakdown of ratings for the three sub-criteria. The data shows that 87% of the projects rated against the sub-criterion on communication between finance and project management achieved S/HS performance, whereas the proportion of projects rated S/HS against adherence to financial policies and procedures, and completeness of financial information and was 62% and 61% respectively.

Figure 21. Financial management (overall), 2014-2023

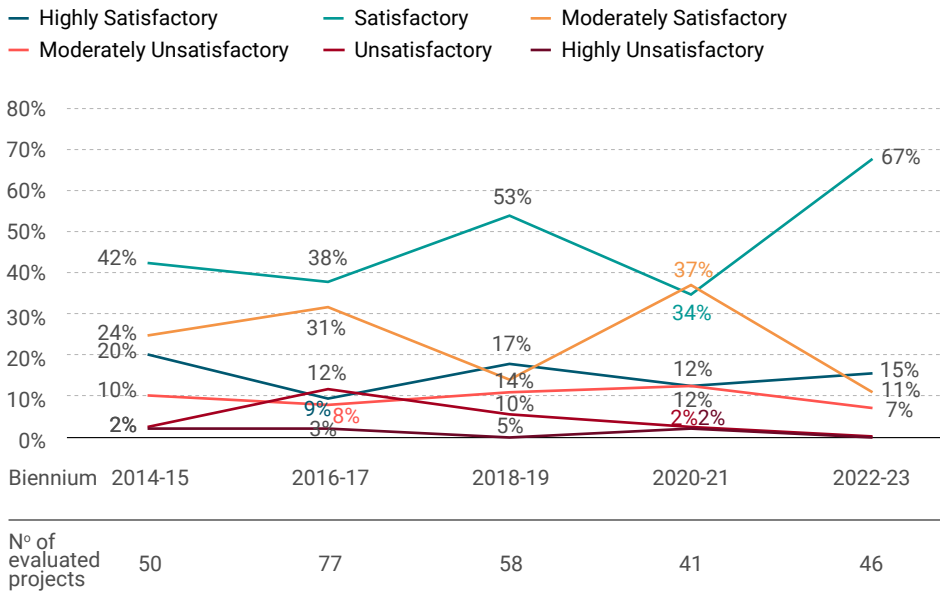
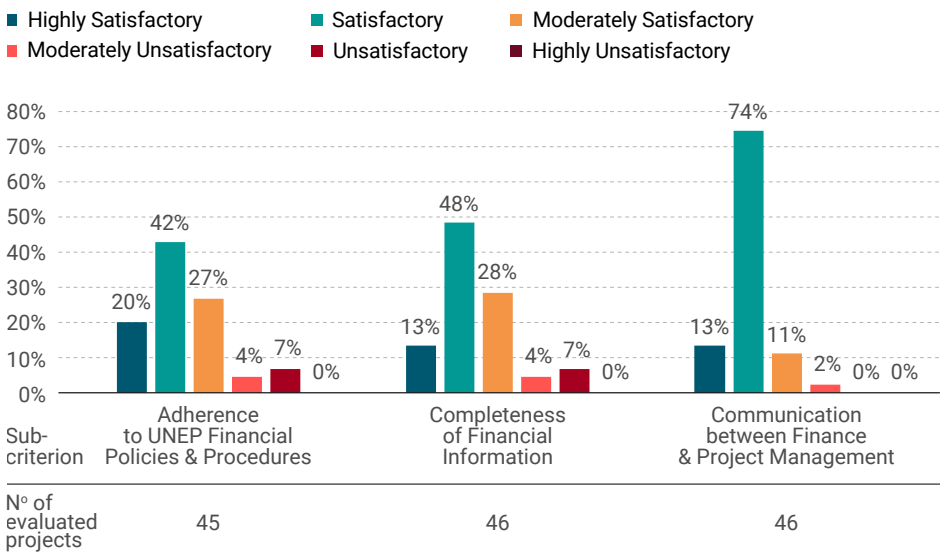


Figure 22. Financial management by sub-criteria, 2022-2023



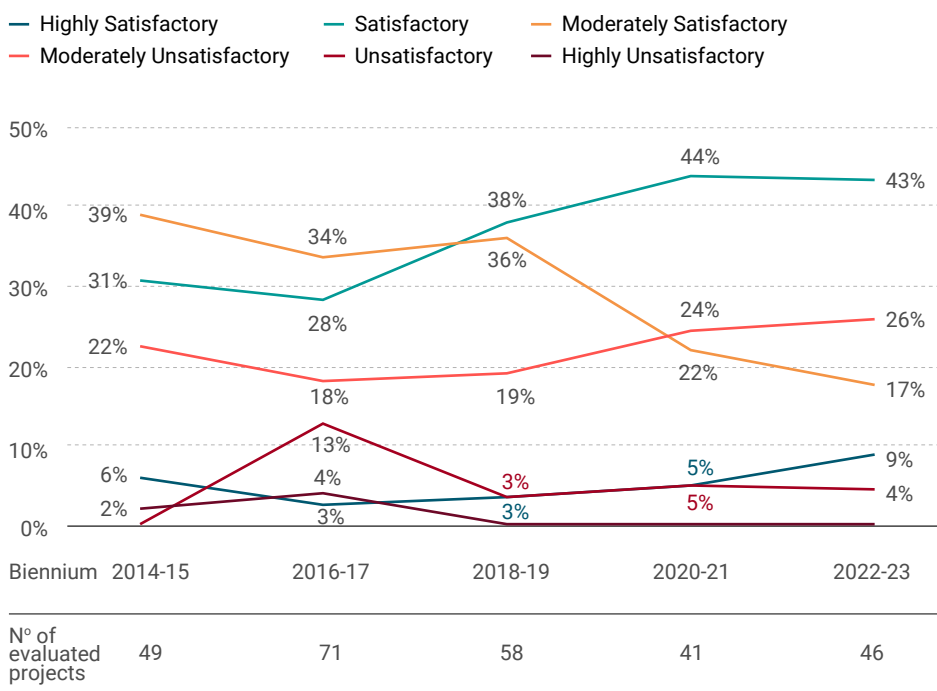
An issue that continues to arise in evaluations, and was highlighted in the previous biennium, is the lack of institutional guidance in defining, estimating, reporting, and verifying co-finance (cash and in-kind contributions). This continues to be the case especially among projects funded by the GEF.

Experiences suggest a need for better protocols on managing project co-finance due to insufficient clarity on the definitions, methods for estimation (at design and implementation), and procedures for monitoring and reporting of co-financing. The differentiation between cash and in-kind co-finance is also important as both need to be accounted for. While evidence of cash co-finance is often available through financial transaction documents, in-kind co-finance is often more complex and requires further guidance in terms of definitions, means of valuation, recording provision and validation.

Monitoring and reporting

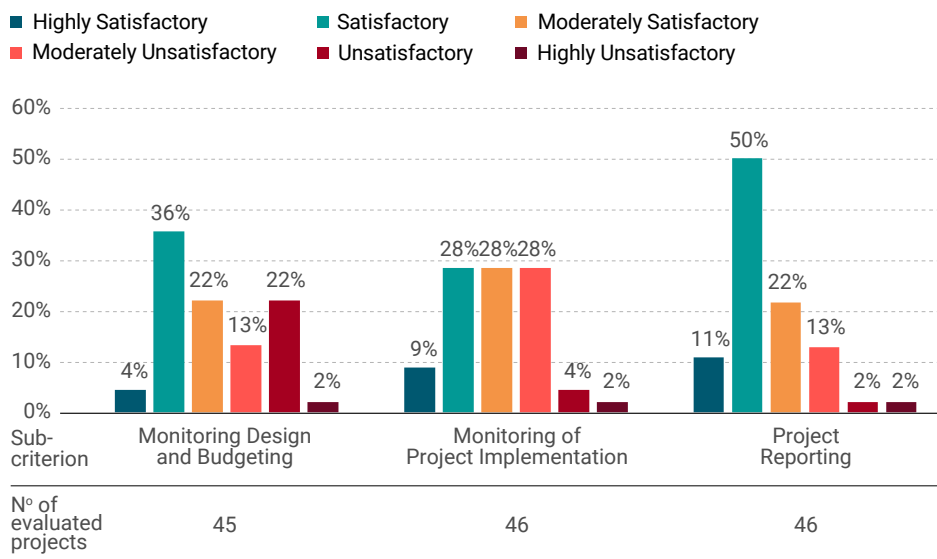
Evaluations assess monitoring and reporting across three sub-criteria: monitoring design and budgeting, monitoring implementation and project reporting. Figure 23 below shows the general trend in performance for the overall monitoring criterion over a 10-year period (2014-23). In 2022 and 2023, 52% of the evaluated projects achieved S/HS ratings, compared to 49% of projects within the same performance range in the previous biennium (2020-21) and an average over the previous four biennia of 40%. However, up to 30% of the projects evaluated against this criterion in 2023-23 were assessed as being either 'Unsatisfactory' (4%) or 'Moderately Unsatisfactory' (26%).

Figure 23. Monitoring and reporting, 2014-2023



The performance of the three sub-criteria on monitoring and reporting are shown in Figure 24.

Figure 24. Monitoring and reporting by sub-criteria, 2022-2023



Monitoring design and budgeting

Ideally, a project should be supported by a sound monitoring plan that is designed to track progress against specified results including at disaggregated levels by gender, vulnerability or marginalisation. Evaluations assess the quality of the design of the monitoring plan, the funds allocated for its implementation, the suitability (relevance and appropriateness) of indicators (both for outputs and outcomes), as well as the methods used for tracking progress against them as part of a conscious results-based management. In 2022 and 2023, 40% of the projects achieved S/HS ratings for having a sound monitoring plan and associated budget.

Monitoring of project implementation

The extent to which the monitoring system was operational and facilitated the timely tracking of results, and progress towards projects objectives, throughout the project implementation period is assessed under this sub-criterion. Evaluations also consider the quality of the information generated by the monitoring system during project implementation, and how it was used to adapt and improve project execution, and achievement of outcomes and their sustainability. In 2022 and 2023, 37% of the projects achieved S/HS ratings for tracking of their results.

Project reporting

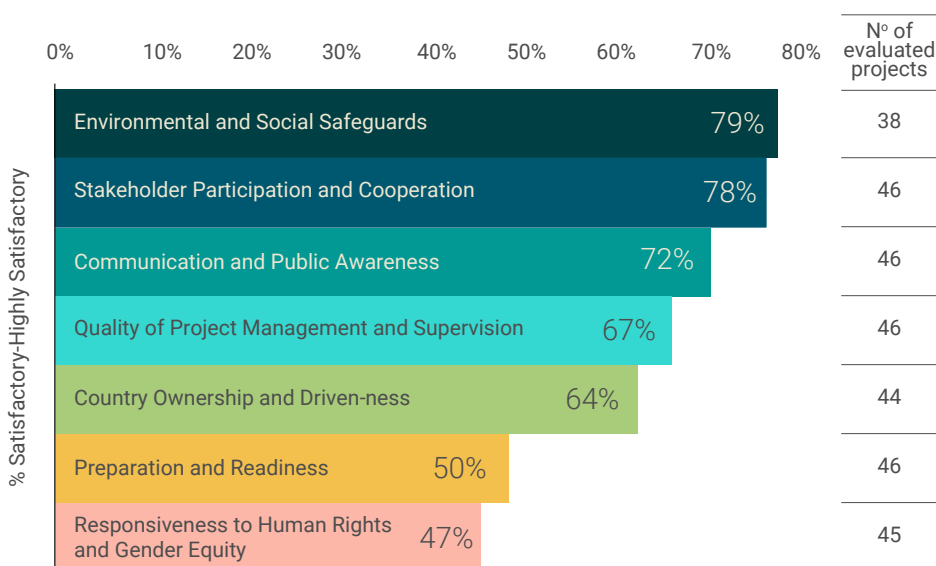
Under this sub-criterion, projects are assessed to determine the extent to which both UNEP and donor reporting commitments have been fulfilled. In 2022 and 2023, 61% of the projects evaluated against this criterion were rated S/HS. One of the key areas of weakness that continues to be observed, especially in projects not attaining the S/HS level, is the lack of distinction made between reporting on project implementation (what has happened) and monitoring implementation against the expected results. In this regard, greater awareness that monitoring should routinely provide a comparison between what has happened and what was planned, is needed. In 2022 and 2023, 61% of the projects achieved S/HS ratings for the quality of reporting.

Reporting, like monitoring, is a core responsibility of the project management team. Monitoring and reporting of project progress against both output- and outcome-level results plays a critical role in informing management of the need to make corrective action during implementation.

In summary, approximately half of the projects evaluated in the last two biennia have not reached the desired performance target (HS/S) for the criterion on monitoring and reporting. **Since monitoring and reporting are aspects of project management that are fully under UNEP's control or oversight, this is an area for continued management attention.**

Factors affecting project performance

Figure 25. Projects rated 'Satisfactory' and better in the factors affecting performance, 2022-2023



This section describes the assessment of cross-cutting themes that encompass factors that commonly affect project performance. These factors include: preparation and readiness; quality of project management and supervision; stakeholders’ participation and cooperation; responsiveness to human rights and gender equity; country ownership and driven-ness; and communication and public awareness.

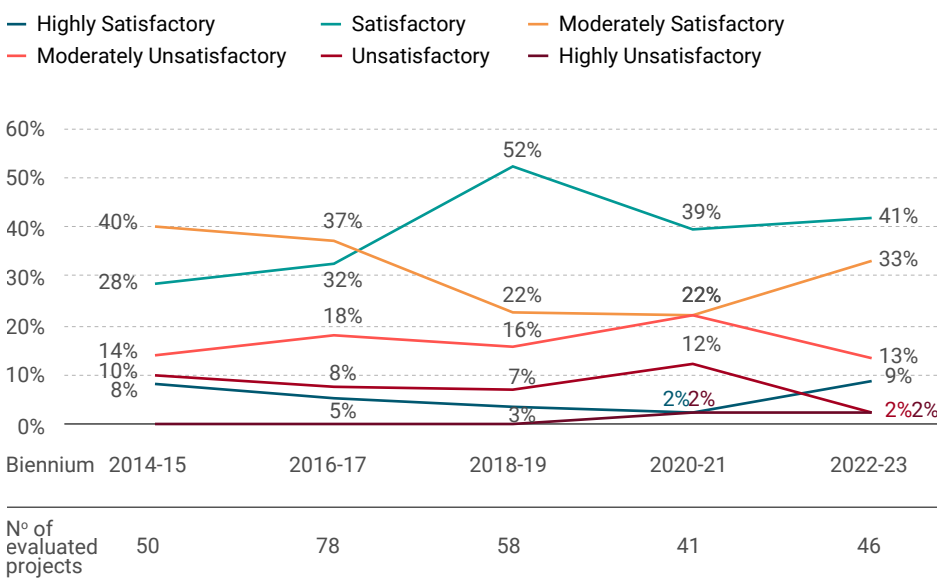
Project performance against these factors for evaluations completed in 2022 and 2023 are presented in **Figure 25** above. Learning from recommendations and best practice examples from evaluations are included in the analyses presented.

Preparation and readiness

This criterion focuses on the inception or mobilisation stage of the project (i.e. the time between project approval and first disbursement). Evaluations assess among other things, the measures taken to address weaknesses in the project design or changes occurring at project mobilisation, engagement with stakeholder groups, confirmation of partner capacity, development of partnership agreements, initial staffing and financing arrangements.

As illustrated in **Figure 26** below, in 2022 and 2023 only half of the evaluated projects (50%) achieved a rating of S/HS for the measures taken prior to project inception. The average performance against this criterion across the previous four biennia was 42% (S/HS rates were 36% in 2014-15; 37% in 2016-17; 55% in 2018-19; and 41% in 2020-21).

Figure 26. Preparation and readiness, 2014-2023



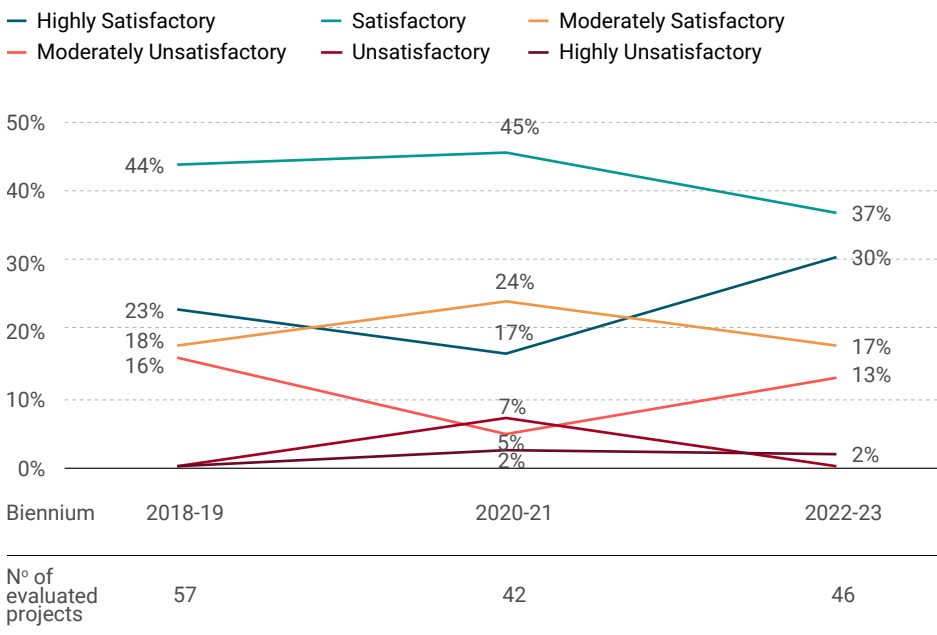
From the experiences recorded in evaluation reports, lengthy delays at inception are often due to the amount of time required to: sign MOU/agreements, disburse funds, and recruit project staff at partner level. The inception workshop is critical for confirming any adjustments to the project design, results and/or partnerships, and this is particularly important where the implementing countries/partners are either not known during project design or have been changed after project approval.

Quality of project management and supervision

Under this criterion, the effectiveness of project management is assessed with regard to: provision of leadership towards achieving the planned outcomes; managing team structures; maintaining productive partner relationships (including steering groups, etc.); maintaining project relevance within changing external and strategic contexts; communication and collaboration with UNEP colleagues; risk management; use of problem-solving; project adaptation and adaptive management; and overall project execution.

As the sub-criteria under Factors Affecting Performance were adjusted in 2017, data for this criterion only exist from 2018 onwards. The data in Figure 27 below show S/HS ratings have remained quite similar over time at a level of 67% in the current biennium, 67% in 2018-19 and 62% in 2020-21.

Figure 27. Quality of project management and supervision, 2018-2023



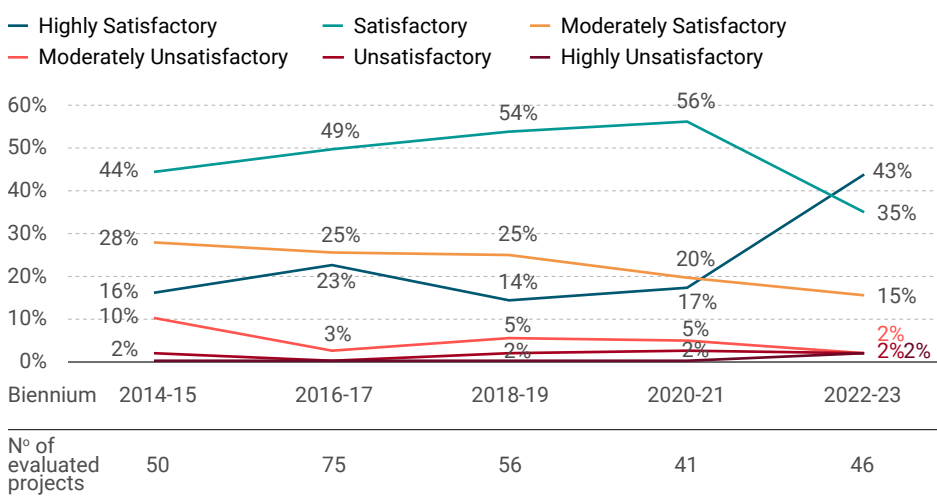
Based on learning from past evaluations, it is very important to strengthen the institutionalisation of project documentation. Project implementation periods are typically 4-7 years, and many receive extensions. Consequently, staff turnover in the project management role is not uncommon as individuals pursue their own career development. When staff leave or move, project information should not be leaving with them; there needs to be routine institutional retention of critical project information to allow for a smooth transition in the project management and or oversight role. In addition, the value of supervision missions should also not be underestimated; they should be encouraged, clearly recorded, and with written justifications when they do not take place.

Stakeholders' participation and cooperation

The assessment against this criterion considers the quality and effectiveness of all forms of communication and consultation with stakeholders²¹ throughout the project life, and the support given to maximise collaboration and coherence between various stakeholders, including sharing plans, pooling resources and exchanging learning and expertise. The inclusion and participation of all differentiated groups, including gender groups is also considered.

This is one of the criteria against which projects have had substantial proportions of S/HS ratings over the years. As shown in **Figure 28** below, the proportion of projects that were rated S/HS against this criterion in the 2022-23 biennium was 78%, higher than the average of 68% over the last four biennia (60% in 2014-15; 72% in 2016-17; 68% in 2018-19; and 73% in 2020-21).

Figure 28. Stakeholders' participation and cooperation, 2014-2023



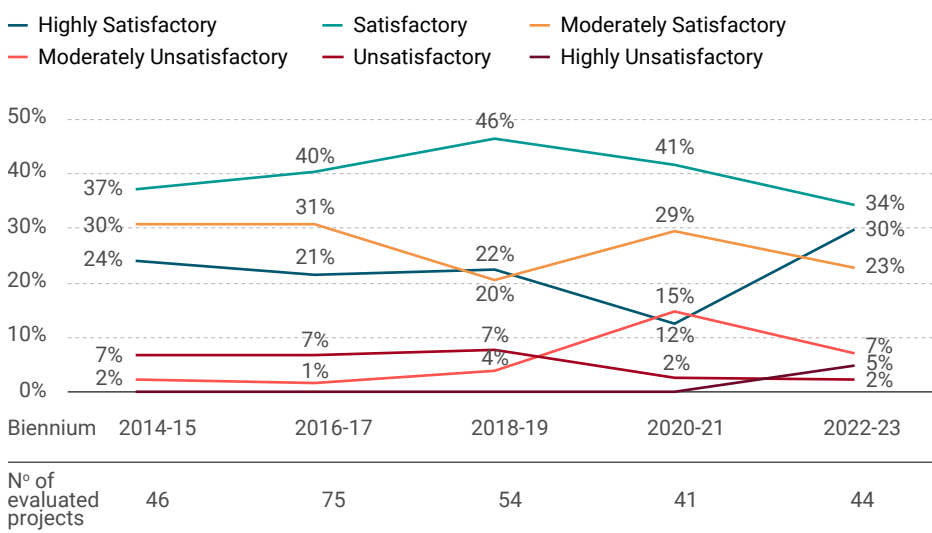
²¹ The term 'stakeholder' should be considered here in a broad sense, encompassing all project partners, duty bearers with a role in delivering project Outputs and target users of project Outputs and any other collaborating agents external to UNEP.

The main learning related to this criterion, based on past project evaluations, is that the absence of critical stakeholders during the project lifetime has a significantly adverse effect on project implementation and sustainability of results. It is important to ensure that the stakeholders who should be engaged in the project (i.e. in terms of interest, commitment, knowledge, decision power, as well as those who should benefit) are selected on a targeted basis, and that the selection criteria are documented and shared. This includes the selection criteria for demonstration/implementation sites.

Country ownership and driven-ness

This factor is concerned with the level of ownership generated by the project over outputs and outcomes, and that is necessary for long-term impact to be realised. Evaluations assess the quality and degree of engagement of government/public sector agencies in projects. They consider the involvement not only of those directly involved in project execution and those participating in technical or leadership groups, but also those official representatives whose cooperation is needed for change to be embedded in their respective institutions and offices (e.g. representatives from multiple sectors or relevant ministries beyond Ministry of Environment).

Figure 29. Country ownership and driven-ness, 2014-2023

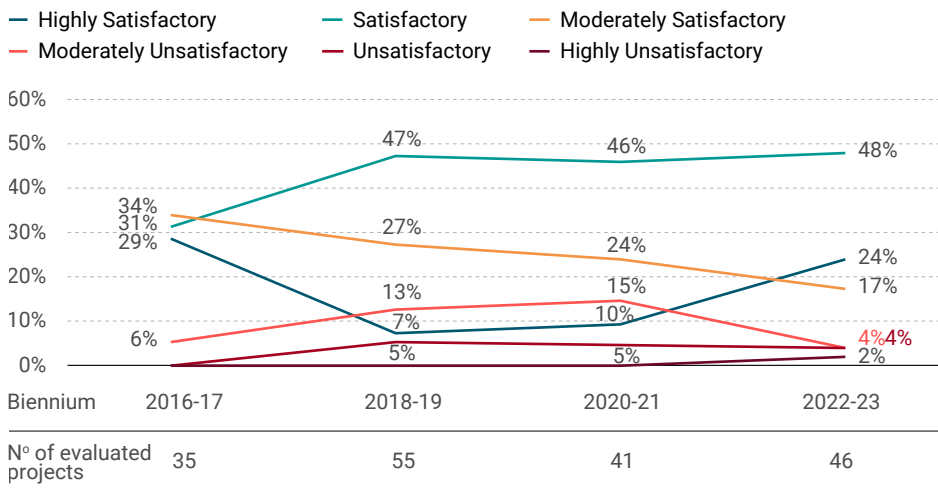


Project performance against this criterion in 2022 and 2023 shows that 64% achieved S/HS ratings, indicating a higher proportion than in the last biennium (2020-21) with 53% projects in the same performance range. Nearly one third of projects (30%) achieved an HS rating in 2022 and 2023, the highest percentage over the last four biennia for which the average was 60% (see Figure 29 above).

Communication and public awareness

This factor was introduced in 2016 to assess the effectiveness of: (i) communication of learning arising from the project during its life, and experience-sharing between project partners and interested groups; and (ii) public awareness activities that were undertaken during the implementation of the project to influence attitudes or shape behaviour among wider communities and civil society at large.

Figure 30. Communication and public awareness, 2016-2023



The data in **Figure 30** above show that 72% of the projects evaluated against this criterion in 2022 and 2023 performed at the S/HS level. As can be seen from the data, the proportion of S/HS rated projects in the 2022-23 is higher than that average of 56% across the previous three biennia (60% in 2016-17; 54% in 2018-19; and 56% in 2020-21).

For projects that have performed at S/HS level, it may be said that the communication channels and networks, as well as feedback channels, were used effectively, including meeting the differentiated needs of gendered or marginalised groups.

From experiences reported in evaluations, projects with sub-optimal performance against this criterion include those where communication/knowledge-sharing products were produced in the latter stages of project implementation and could then not play the catalytic/awareness-raising role intended because they were not available to the intended audiences in a timely manner. The realistic timing of communication products within the project implementation timeframe requires more attention, and their role within the Theory of Change adjusted accordingly. In addition, projects should integrate methodical communication strategies with specific budget allocations made available for their implementation.

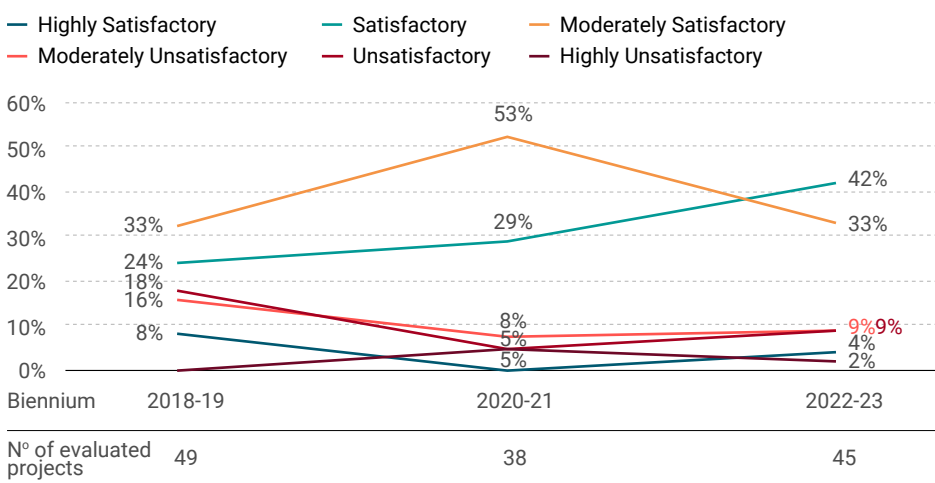
Responsiveness to human rights and gender equality

The criterion on responsiveness to human rights and gender equality was introduced as a criterion under the Factors Affecting Performance in 2018, therefore the data collected on this criterion covers only three biennia.

Evaluations assess the extent to which projects have taken into consideration: (i) possible inequalities (especially those related to gender) in access to, and the control over, natural resources; (ii) specific vulnerabilities of disadvantaged groups (especially women, children and those living with disabilities) to environmental degradation or disasters; and (iii) the role of disadvantaged groups in mitigating or adapting to environmental changes and engaging in environmental protection and rehabilitation.

As shown in **Figure 31** below, on average 35% of the projects assessed for this criterion over the three biennia (2018-19, 2020-21, and 2022-23) received a S/HS rating. To date, 2022 and 2023 have had the highest record of projects achieving S/HS ratings (46%)²².

Figure 31. Responsiveness to human rights and gender equality, 2018-2023



Learning from completed evaluations has shown that being responsive to human rights and gender may not involve additional costed activities and may, therefore, not be very clear in the workplan or budget. The integration of gender and human rights in project implementation therefore needs a conscious effort and action to be included in a project's design including the Theory of Change²³.

²² See Annex 2 for a list of evaluated projects and their rating against the Human Rights and Gender Equality criterion.

²³ The Evaluation Office is already requiring consultants to incorporate an 'equality' Driver/Assumption in the Theory of Change produced during an evaluation process.

UNEP takes part in the UN Sector-wide Approach on Gender Equality and the Empowerment of Women (UN-SWAP) and, as part of this process, an external consultant is commissioned by the Evaluation Office to review all the project evaluations completed within the year. Although this exercise primarily focuses on how well the evaluation function is addressing gender equality in its evaluation processes and approaches, the exercise also provides some insight into UNEPs performance at project level. A summary of the findings from the UN-SWAP Report on UNEP is presented in **section 8.2** of this report.

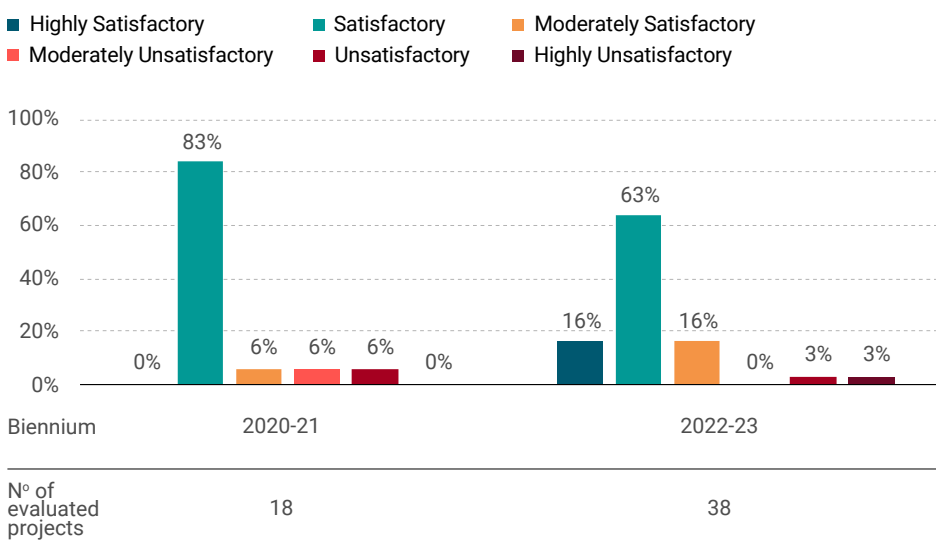
Environmental and social safeguards

This factor was included for the first time in the 2020-2021 biennium, therefore the data collected on this criterion covers only two biennia with 56 projects being evaluated against their responsiveness to addressing environmental and social safeguard risks.

Under this criterion evaluators consider whether: UNEP requirements were adequately met for screening environmental and social safeguards; environmental and social risk ratings were reviewed on a regular basis; project implementation was monitored for possible safeguard risks; responses (where relevant) to safeguard issues were adequate; and reporting on the implementation of safeguards risk management was undertaken.

Performance under this sub-criterion has high proportions of S/HS ratings in both biennia to-date (2020-21 and 2022-23), with 79% of the projects assessed attaining S/HS ratings in 2022 and 2023 (see **Figure 32** below).

Figure 32. Environmental and social safeguards, 2020-2023



Chapter 4

Key Findings from Strategic Evaluations

This chapter presents abridged evaluation findings from four evaluations of high strategic importance that were completed in 2022 and 2023.

Evaluation of the UNEP subprogramme on environmental governance, 2014-2021

The subprogramme on environmental governance (SP-EG) was established in 2009 in response to a call from the UNEP Governing Council to “set a normative agenda for international environmental governance (IEG) and a functional IEG system that provides the international framework to support governments in successfully addressing environmental challenges and meeting their commitments and as a precondition to carry out other UNEP activities effectively.” The current Medium-Term Strategy 2022-2025 explicitly places the SP-EG as a foundational programme that contributes to activities and

results of other subprogrammes and that aims to support the three thematic subprogrammes.

This evaluation assessed the SP-EG over the period 2014-2021²⁴. The evaluation scope included all projects and programmes under the SP-EG Programme of Work (PoW) for each year in the evaluation period, including consideration of the forward PoW for 2022-2023. The evaluation did not cover the work of Multilateral Environmental Agreement (MEA) Secretariats but did cover SP-EG activities that supported the work of MEAs such as MEA secretariat support, joint projects, support for Conference of the Parties (COP)-preparations and trainings for negotiators.

Achievements

The SP-EG has consistently targeted three areas over the evaluation period: i) a normative approach focused on preparing and developing laws and standards; ii) supporting and striving for coherence of global environmental governance with MEAs, the Sustainable Development Goals (SDGs) and the United Nations system; and iii) supporting active environmental governance capacity building within UNEP, more broadly within the UN system and also across Member States.

At a global level, SP-EG responds to an increasing focus on environmental governance issues. At the strategic level, the SP-EG is intrinsically linked to UNEP's other subprogrammes with governance instruments identified as fundamental to achieving other environmental goals. However, strategic documents, such as the MTS, do not include a clear definition of environmental governance (EG) or practical guidance for programming or coordination that can shape the approach across the organization. As a result, there was variance in the understanding of environmental governance across UNEP and difficulty in articulating the added value of EG activities.

The SP-EG portfolio was found to be fragmented and not reflect a clear strategic direction or pathway towards articulated goals. The portfolio did respond to United Nations Environment Assembly (UNEA) and funding partner priorities; however, the relevance at the national level for individual Member States was less clear. The portfolio, as a whole, did not operate on systematic country needs assessments or prioritization although there were some positive approaches by some individual projects and programmes. Nonetheless, specialized environmental law expertise, regional environmental governance activities and support to MEAs were considered highly relevant as a means for achieving EG-related results.

²⁴ Link to report: <https://wedocs.unep.org/handle/20.500.11822/41594>

Most projects have demonstrated achievement of positive results that contribute to improved EG Outcomes. Three key categories of activities contributing to these higher-level Outcomes have been identified, namely, i) legal support, ii) MEA support, and iii) EG capacity building to enable direct benefits in environmental governance.

The work under the Montevideo Programme, as part of the legal support category, over the years has generated spin-off effects contributing to the legal aspects of international agreements on mercury and the management of transboundary freshwater resources as well as training programmes.

There was insufficient synergy and some duplication of effort between subprogrammes. This has led to a dilution of EG-related activities across UNEP subprogrammes and monitoring of results and progress towards expected accomplishments for EG is difficult. Of particular importance, was the link between the SP-EG and the subprogramme on Science-Policy (previously the Environment Under Review SP) and the essential link between these two subprogrammes was not adequately acknowledged, fostered or supported by institutional processes. The linkages to other subprogrammes expected of the SP-EG (especially going forward as a foundational subprogramme) was not adequately reflected in organizational and strategic tools such as Programmes of Work or Results Frameworks. Further, current management structures and processes focus mainly on the divisions and branch activities through individual projects or programmes, rather than results of the subprogramme as a whole.

Sustainability

There were good practice examples of sustainability within the SP-EG portfolio of projects where there was evidence of institutional change and capacity built. However, common threats to sustainability across the subprogramme included limited systematic identification of and responsiveness to country priorities and insufficient engagement of decision-makers to generate ownership for longer-term results.

The SP-EG has not generated strong evidence to confirm contribution to mainstreaming of human rights, gender and considerations of vulnerable populations into SP-EG activities, although some positive progress towards this is evident. The work of the SP-EG officers to support efforts within the Human Rights Council and UNGA to adopt a human right to the environment has a significant impact potential worldwide.

Lessons learned

There is limited shared knowledge and understanding more broadly within UNEP about EG. As such the value of the subprogramme is somewhat undermined and

there are insufficient guidelines and tools to set specific goals, identify intended results or prioritize activities accordingly. This contributes to a fragmented approach to the design and implementation of the subprogramme.

The support to MEA priorities is of significant importance for global and national EG. The work of the SP-EG in supporting common priorities of MEAs such as information management through InfoMEA, collaborations through the MEAs focal point network and the MEAs gender network and the coordination work through the regional offices is highly valued but under-staffed.

The approach to programming expected through the sub-programme modality framed around themes (e.g. environmental governance, climate action, etc.) is not reflected in current management structures which continue to be driven by divisional processes and structure focused on individual projects. Changes to these structures, including creating 13 programmes, situating the SP Coordinators in the Policy and Programme Division and developing a new 'delivery model' aim to address these challenges but are likely to take time to have a real impact.

There is a huge potential for involvement in the UN Reform process to increase with several benefits for UNEP and the SP-EG. Firstly, enhanced engagement of the SP-EG in common country programming processes would allow for increased knowledge and understating about contextual factors and priorities.

Recommendations

Recommendation 1: UNEP's leadership role and approach to environmental governance needs a clearer focus, strategic priorities and more specific Outcomes. The SP-EG should develop, in collaboration with other subprogrammes, Divisions and MEAs, an improved overall strategy and mechanisms to strengthen its approach to EG across UNEP in line with the MTS 2022-2025 UNEP thematic priorities. This needs to include a clear, working definition of EG to inform the SP-EG's scope and activities. The SP-EG needs a targeted results framework (especially immediate Outcomes, direct indicators and units of measure) to help guide the SP-EG, as foundational and cross-cutting subprogramme, in relation to the UNEP thematic priorities. Adopting a SP-EG Theory of Change (ToC), nested with the three thematic subprogramme ToCs, would aid design and implementation.

Recommendation 2: Environmental Governance requires a more cross-cutting and mainstreamed approach to developing its PoW to achieve the MTS 2022-2025 vision of SP-EG as a foundational subprogramme. The SP-EG should undertake a strategic review across UNEP to improve programming. This would include understanding what activities of relevance to EG are being pursued in other subprogrammes and where synergies may be best realized. The actions arising

need to be applied through an integrated PoW with stronger links with other subprogrammes to capitalize on opportunities to strengthen EG collaboratively across subprogrammes and on cross-cutting governance issues such as gender and human rights. The SP-EG workplan should be linked with and demonstrate the added value of the SP-EG to the three UNEP thematic priorities (climate action, chemicals/pollution action and nature action) across all divisions.

Recommendation 3: Environmental Governance key functions require an amended structure, linked to the Theory of Change and Outcomes identified in Recommendation 1. Clearer distinction is required between the three functions of EG, (i) Legal support, (ii) MEA support and (iii) strengthening EG capacity more broadly within the UNEP, the UN and Members States, is required so that these functions can be more clearly addressed within the portfolio and through collaboration.

Recommendation 4: The SP-EG needs to be proactive in its supporting role across UNEP and clarify and seek the specific expertise required to fill roles in support of cooperation, skills enhancement activities and establishing specific coordination mechanisms. Closer cooperation between the subprogrammes, MEAs and partners in EG capacity strengthening on the development of joint indicators, units of measure for coordination and support would enhance cooperation and Outcomes.

Recommendation 5: UNEP should make more of the opportunity of UN Reform to mainstream EG initiatives as well as other UNEP activities at the national level. This requires a more active initiative to engage with key countries in the CCA and United Nations Sustainable Development Cooperation Framework (UNSDCF) to leverage activities in EG and build learning to support future UNCT, MEA and Member State initiatives. Clearer, more systematic pathways for identifying country priorities for EG to link with the UNCT processes for CCA and UNSDCF would provide a substantial opportunity to enhance the mainstreaming of environmental governance to the entire UN programme of support across each country.

Independent evaluation of UNEP Policy and Strategy for Gender Equality and the Environment, 2015-2020

UNEP has had three gender policies or strategies. The first 'UNEP Gender Strategy' was developed in 2000 and responded to decision 20/9 of 5 February 1999 on the role of women and the environment and development. The second UNEP 'Gender Plan of Action' (GePA) was published in September 2006. It contained a framework for integrating a gender perspective within UNEP and in its programmes for the period 2006-10. The third 'UNEP Policy and Strategy for

Gender Equality and the Environment' was adopted in 2015 and developed as an adjunct of UNEP's Medium-Term Strategy 2014-2017.

The 2015 GPS provided UNEP with a framework for addressing gender equality and women's empowerment in a coherent and coordinated manner in its policies and programmes. It was intended to be implemented over a period of four years (i.e., 2014-17). It is understood that the 2015 GPS continues to direct UNEP's gender response until a new strategy is published.

This was the first independent evaluation of the implementation of UNEP's Policy and Strategy for gender mainstreaming during the period from 2015 to 2020 based on the 2015 GPS²⁵. The purpose of the evaluation was to assess UNEP's progress in mainstreaming gender at the institutional level and throughout UNEP activities, and UNEP's programmatic efforts in promoting GEWE in terms of access to, and control over, environmental resources.

Achievements

UNEP's GPS contained promising statements and ambitious objectives, but the absence of a gender focus in UNEP's main strategic framework, the Medium-Term Strategy (MTS) 2014-2017 and 2018 -2021, limited the integration of gender in the organization's programmatic work as well as in operationalizing its business model.

The absence of a results-based framework with expected accomplishments, indicators (in the MTS) and also the lack a reporting system with gender results, was a major obstacle in terms of identifying gender results that can be attributed to the impact of the GPS.

UNEP's performance in relation to the UN system-wide Action Plan (UN-SWAP) during the period 2015-2020 has also been low (47% in 2019 and 2020).

The establishment and implementation of the Gender Marker has contributed to strengthen the integration of gender in project proposals but has not translated into the systematic mainstreaming of gender in UNEP's programmatic work.

UNEP's most notable success has been promoting gender parity, particularly within the senior professional staff levels of the organization.

The lack of a tracking mechanism has prevented UNEP from quantifying its investments in gender mainstreaming and weakened accountability for the implementation of the GPS.

²⁵ Link to report: <https://wedocs.unep.org/handle/20.500.11822/40571>

The GSU and the gender focal points (GFPs) have contributed to advancing gender mainstreaming in UNEP's projects, but the support of the gender architecture to UNEP's mandate on GEWE has not been consistent or systematic over the period 2015-2020. The absence of results statements and indicators in the GPS and implementation plan has resulted in a lack of clarity as to the role of the GSU and the role of GFPs as well as to the responsibilities of staff in achieving GEWE in UNEP's work.

The inconsistent integration of gender, the only exception being GEF funded projects, in project log frames and budgets has prevented the current monitoring system from effectively capturing gender results. As a result, gender information has not been systematically collected and used to support decision making.

Systematic sharing of lessons learned on GEWE for use by programme designers, managers and evaluators could help improve programme design, implementation, monitoring and evaluation across UNEP programming.

Sustainability

The GPS has promoted visibility of gender mainstreaming in key publications, including the 2016 flagship "Global Gender and Environment Outlook" (GGEO). Since 2007, the GSU has been involved in reviewing a significant number of publications produced by UNEP to ensure the integration of gender.

Key lessons

Though it played a critical step to introduce gender mainstreaming in UNEP, the GPS has become largely symbolic of UNEP's unrealised ambitions in gender. The overall lesson learned is that without senior management attention, resource allocation, results frameworks, monitoring systems, a functioning gender architecture, and a concerted effort at sensitizing and capacitating staff as to what mainstreaming gender entails within UNEP programmes, gender quickly became a forgotten priority.

Periodic reporting to the governing body on the nature and type of gender programming is a crucial indicator of agency commitment. Reporting on gender parity advances and isolated examples of programmes working on gender in a few paragraphs in annual progress reports is not a substitute for systematic reporting to UNEA on an annual or biennial basis.

The Gender and Safeguards Unit was the central hub responsible to spearhead the implementation of the GPS, develop training and capacity building, report to donors and the UN-SWAP, clear publications, organize key outreach efforts and introduce and manage the Gender Marker system. Without sufficient capacity

and an annually resourced plan of action which was monitored by the SMT, the GSU was unable to meet the ambitious expectations of the GPS.

Reporting to the UN-SWAP on an annual basis has limited effectiveness unless it is seen as an instrument of relevance to headquarters, regional and country-level managers and involved annual internal stock-taking of the type of gender mainstreaming taking place within the organization.

The most consistent gender messaging during 2015-2020 was in terms of Gender Parity which speaks to the impact of senior level attention to this aspect of the GPS. This not been matched with attention to gender mainstreaming in UNEP's programme and outreach efforts.

Key recommendations

Strategic relevance

Appoint a high-level, internal Task Force to undertake a time-bound exercise to foster organization-wide attention to gender mainstreaming in both programme and institutional domains. This exercise should be led by a high-profile champion such as the Deputy Executive Director to devise an updated strategy involving a rethinking of UNEP's approach to Gender with particular focus on the opportunities presented by UN Reform efforts at the country and regional levels, and with the GSU serving as secretariat. This new thinking could involve revisiting the Theories of Change (TOC) in the Programmes of Work (PoW) to ensure that gender is incorporated as a critical driver to inclusive change and in indicator frameworks.

Annual UN-SWAP reporting should become an instrument for managers to periodically stock-take, report and plan for improvements so that reaching the MTS goal of an 82% aggregate score in the UN-SWAP self-assessment process is seen as an institutional challenge for all staff, UNEP wide.

Appoint an External Gender Expert Advisory Group: The Executive Director should consider convening a group of independent external experts and gender champions to provide the SMT and the Task Force with guidance on key areas relating to the nexus between gender and the environment.

Effectiveness

Use UN-SWAP as an accountability and management instrument: The UN-SWAP methodology is designed to actively involve all key functions of an organization to keep track and report on efforts to mainstream gender on an annual basis. To make optimum use of this instrument, the SMT should require all key managers

to become familiar with the UN-SWAP methodology and actively engage in formulating their individual responses and rating their progress.

Optimize the application of the Gender Marker: The GSU should ensure that staff, in particular subprogramme Coordinators, project managers/implementers are familiar with the Gender Marker and understand UNEP commitments in this area. Gender equality results that are to be achieved by the project, should be incorporated at the design stage of projects.

Formulate a Communication Strategy to ensure transparency in initiatives vis-a-vis Gender Parity: Senior Management should take steps to ensure transparent communication on appointees such as sharing the credentials of the women selected for senior posts (which would help decrease a perception that women were being selected because of the pressure to achieve parity). Furthermore, steps should be taken to address and report on gender parity in the General Service Staff cadre of staff.

Financial management

Establish financial benchmarks and a tracking system for gender allocations linked to Gender Marker: The SMT, in coordination with the GSU, should also establish a financial benchmark for resources spent on GEWE, including at the sub-programme level. The financial benchmark should combine financial contributions to address specific gender equality issues, with contributions towards the integration of gender in other environmental issues. Additionally, it should separately track UNEP's own expenditure and UNEP's transfers to partners for project implementation.

The Corporate Services Division should put in place a financial tracking system to quantify funds allocated for the promotion of GEWE. This tracking system should be linked to the Gender Marker to systematically track gender related allocations across projects with explicit gender components.

Consult and share experiences with other UN agencies: To establish the financial goals, UNEP should consult other UN agencies with experience in this area, including UNICEF, UNDP, UNFPA, ILO, WTO and WMO.

Efficiency

Strengthen the Gender Architecture: UNEP should strengthen the gender architecture which involves appointing Gender Advisers at the sub-programme level, with a clear role for the GSU and a revitalizing of the Gender Focal Point function as a strong network with the objective of setting up a community of practice on gender.

Monitor and report on impact in terms of gender results: Once the new Integrating Planning, Management and Reporting System (IPMR) is fully operational, Programme Performance Support Unit (PPSU), Policy and Programme Division with support from the GSU should ensure that monitoring and reporting move away from describing activities and compiling statistics and is tied to higher level gender results.

Update the Guidance Note on Gender and integrate it as part of the overall guidance to evaluation consultants: The Evaluation Office should compile a compendium of lessons learned on gender interventions that help describe the six-point satisfaction scale of the Evaluation Matrix.

Report on gender quality impact in UNEP projects (as stipulated in PoW 2022-2023 under Executive Management “number of UNEP projects assessed for gender quality impact” as an indicator of the degree of implementation of Resolution 4/17 on promoting gender equality and human rights and empowerment of women and girls in environmental governance) should go beyond numbers of evaluations and consider providing fuller accounts of gender results captured in evaluations for use in lesson learning and reporting. These could be showcased in the annual Progress Reports and presented to the SMT and CPR as appropriate.

Revisit the use of gender goals in e-PAS. UNEP should review the guidance to incorporate gender goals at the e-PAS level and identify relevant tasks for the year at hand for each staff member that will help strengthen gender mainstreaming.

Terminal Evaluation of the UNEP Inquiry into the design of a sustainable financial system and Interim Review of the UNEP/GEF Aligning the financial system and infrastructure and infrastructure investments with sustainable development (Components 1 and 2)

The past ten years have seen major gains in the greening of the global financial system. Spanning most of that decade, the projects assessed here made important contributions to this effort. However, the scale of sustainable finance across and within countries remains too small, and the velocity of its growth too slow, to fully address the requirements of the Sustainable Development Goals and the Paris Accord in the real economy. In addition to fulfilling its accountability function, this evaluation draws lessons from the projects that shed light on what more can be done to accelerate the scaling of sustainable finance and to ensure its impacts make a meaningful and measurable difference to people and the planet.

Approved in 2013 and launched in 2014, the Inquiry Project was completed in December 2021, reporting expenditures of USD 16.25 million. The objective of the project, as stated in the 2013 Project Document, was to ‘identify and develop financial market policy and regulatory options, based on global best practice, which would deliver a step change in their effectiveness in channeling capital to green investments’. The project aim was ‘to engage, inform and guide policy makers, financial market actors and other stakeholders concerned with the health of the financial system and its impacts’.

With its Components 1 and 2 executed by Inquiry, the GEF Project was approved in 2018, began operations in early 2019, and is expected to be completed in December 2022. It has reported direct expenditures for the two components of USD .565 million by early 2022. Addressing greenhouse gas emission mitigation, the GEF Project’s overall objective is to build international consensus to align financial systems with the Sustainable Development Goals and catalyse national regulatory actions and regional sustainable infrastructure investments. Both the Inquiry and GEF projects have contributed to UNEP’s subprogramme on Resource Efficiency, which is primarily operationalized through the Economy Division²⁶.

Two related projects were assessed for this evaluation with performance ratings were presented separately for the Inquiry and GEF projects²⁷.

Achievements

From 2014 through 2018, with ambition and ingenuity, the Inquiry Project embedded policy-change processes for sustainable finance in important global bodies like the G20, G7 and World Bank and in the financial sectors of 20 countries, constituting impressive system-level gains. Inquiry’s core methods included apex-level networking with elite decision-makers, leverage through central banks and financial supervisors, new-knowledge production, strategic communications, external fundraising, and new-platform creation. Since 2019, the GEF Project, with Components 1 and 2 executed by Inquiry, has focused its support on the national roadmaps and next-generation research needed to build sustainable finance markets and policy architecture in six key emerging-market economies, highlighting innovation and facilitating cross-national learning.

Taken as a whole, this trajectory of work received consistent, important support from the UK, China, EU, Germany, Italy, and Switzerland and several philanthropic foundations. Toward the end of the period, the locus of support for Inquiry’s legacy platforms in the G20, insurance, and financial centres shifted to the United

²⁶ Economy Division is now named Industry and Economy Division.

²⁷ Link to report: <https://wedocs.unep.org/handle/20.500.11822/41410>

Nations Development Programme (UNDP), which had expanded its mandate on sustainable finance. For UNEP, two projects—the Green Growth Knowledge Platform (GGKP) and UNEP Finance Initiative (UNEP-FI)—are pursuing follow-on activities through 2022.

From the outset, the strategic intent of the Inquiry Project was to fast-track UNEP’s work on sustainable finance, which yielded important benefits. At the same time, however, Inquiry’s hard-charging, disruptive operating style, its administrative outsourcing to UNOPS, and other factors contributed to institutional tensions and lack of uptake of the project within UNEP. Both the project and the institution bear responsibility for the strains that arose. In addition, the Inquiry Project did not make a concerted effort to address the needs of vulnerable countries and communities, especially small island developing states (SIDS), and very poor countries in general. Nor did the Inquiry Project take a systematic approach to gender equality, although the GEF Project has been a stronger performer in this regard. One unintended consequence of the rise of sustainable finance to which Inquiry and many other organizations may have contributed is the possibility of a higher cost of capital in some emerging economies. Finally, expressly designed for breadth and reach, the Inquiry-GEF set of interventions did not “go deep” into specific verticals, such as the methane, chemicals or plastics industries, or other important sectors where sustainable finance could play a role in accelerating climate-transition efforts.

The Inquiry Project was, overall, rated ‘Highly Satisfactory’. For Components 1 and 2, the GEF Project was, on an interim basis, rated ‘Satisfactory’. Notwithstanding limits in their design and operations, the projects have achieved very real successes and innovations at meaningful scale and developed and refined a productive implementation model.

Sustainability

The Inquiry Project was a very strong performer in terms of effectiveness and sustainability, as well as on stakeholder participation and country ownership. For its part, the GEF Project has to date performed solidly overall, with notable strengths indicated in sustainability and country ownership.

Lessons learned

The convening power of the United Nations remains a powerful asset.

The Inquiry model is available to UNEP and other organizations to be applied to a new set of challenges and opportunities.

Any application of the model should be strictly time-bound, set out clear pathways for exit and follow-up, and perhaps be administered outside UNEP.

Planning, animating, and supporting new, specialized platforms which can be spun off as separate, self-driven entities constitutes an effective strategy for advancing and sustaining the work.

Projects in sustainable finance should collect gender-disaggregated data to promote gender equality; and projects in sustainable finance should document and calculate the monetised value of the in-kind contributions they mobilize from programming partners.

Recommendations

UNEP consider undertaking a comprehensive review of its programming in sustainable finance to renew its role, strategy, and programming instruments with the aims of accelerating the exponential, targeted growth of sustainable investments and ensuring their positive downstream impact, working with UNDP and other partners, especially in vulnerable countries and among marginalized groups. In particular, the sustainable finance findings from the Inquiry project should be used to inform an alignment of private finance and public finance approaches used by UNEP-FI and Economy Division's fiscal team, to create more synergies and strategic mobilization of finance towards achievement of the SDGs under UNEP.

As part of this review, UNEP should assess the extent to which, and through which modalities, UNEP's sustainable finance programming instruments—including the GEF, GGKP, PAGE, UNEP-FI, and other initiatives—should re-set their work with sustainable finance regulators, in coordination with UNDP and other key actors.

UNEP should consider the extent to which elements of an adapted version of the Inquiry model could and should be applied, for a limited period, to a specific programming area or set of issues to raise awareness, engage elite decision-makers, generate new knowledge, facilitate rapid, scaled action, and perhaps incubate and spin-off new, specialized platforms or vehicles.

UNEP should request Economy Division to conduct an analysis on the integration of targeted programming for SIDS, LDCs, indigenous organizations and other marginalized groups across the Division's portfolio. This may lead to a UNEP-wide response led by the Ecosystems Division to which the Economy Division including contributes.

UNEP should request Economy Division to develop and implement a fully-fledged gender equality strategy to be integrated into all its policies, programs, and projects. This may lead to a UNEP-wide response led by the Policy and Programme Division and to which the Economy Division contributes.

UNEP should request the GEF Project, supported by UNEP-FI, to organize an educational webinar with tailored learning materials on sustainable finance roadmaps for members of the V20 and Climate Vulnerable Forum.

Through GGKP, UNEP should strengthen its approach to identifying and meeting co-financing obligations under the GEF Project by systematically and comprehensively documenting and reporting the monetized value of in-kind contributions by its staff and its country and international partners.

Evaluation of the implementation of the Partnership for Action on Green Economy (PAGE) interagency programme ‘Operational Strategy 2016-2020’

Launched by UNEP in 2013 as an interagency programme with a 7-year duration under its Green Economy Initiative, the Partnership for Action on Green Economy (PAGE) was established as a joint programme in early 2014 by its founding agencies: ILO, UNEP, UNIDO, UNITAR, and joined by UNDP in June 2014. Taking up the call for the UN to support countries interested in advancing a green economy and based on PAGE’s early success in supporting Mongolia and Peru in advancing on this trajectory, more countries showed interest to join and more donors came forward with the needed support.

The Operational Strategy 2016-2020 was jointly developed by the five UN agencies. It had a budget of USD 43.5 million funded by the European Commission (EC) and the governments of Finland, Germany, Norway, Republic of Korea, Sweden, and Switzerland. PAGE’s Multi-Partner Trust Fund (MPTF) became operational from January 2017 and ran alongside a donor agreement in which UNEP managed funds provided by the EC.

Under PAGE’s Operational Strategy 2016-2020, support was extended to 20 Partner Countries and Provinces, listed in chronological order of joining PAGE since 2013: Mongolia, Peru, Senegal, Mauritius, Ghana, Burkina Faso, Jiangsu Province (China), South Africa, Barbados, Mato Grosso (Brazil), Kyrgyz Republic, Guyana, Uruguay, Argentina, Guatemala, India, Indonesia, Kazakhstan, Morocco, and Thailand.

The evaluation of PAGE’s Operational Strategy 2016-2020 was triggered by the close of this 5-year strategy implementation involving five UN agencies, each with differing administrative requirements vis-à-vis evaluation. UNEP led this evaluation, which was funded by the EC, on behalf of the five PAGE agencies. In this setting, UNEP’s evaluation procedures were applied²⁸.

²⁸ Link to report: <https://wedocs.unep.org/handle/20.500.11822/42281>

Achievements

The Operational Strategy offered a clear platform for coordination and the pursuit of complementarities and synergies within and beyond PAGE. While governance and management mechanisms were set-up to ensure internal coherence, the Operational Strategy appeared to rely on an assumption of strong national ownership and leadership and did not strongly address PAGE's external coherence (i.e. provide direction as to how PAGE's in-country governance set-up would drive the coordination beyond the inter-ministerial steering committee and avoid overlap and duplication of efforts from other development partners).

Strengths of PAGE's Operational Strategy 2016-2020 were found in its strategic relevance; the coordination and trust developed amongst the five UN partner agencies that succeeded in fostering inter-ministerial dialogue for policy coherence at national level; the role, competence, and approach of the Secretariat in supporting joint programming and implementation and gatekeeping to uphold the PAGE brand; the belief in the importance, value, and practice of peer learning that permeated PAGE's organisational culture; and the introduction of the MPTF that significantly boosted PAGE's efficiency and financial transparency.

PAGE was effective in producing Outputs, particularly in countries where engagement started before the Operational Strategy 2016-2020. However, at outcome level, results achievement was mixed. Outcome 1 (national planning) and Outcome 4 (knowledge base for IGE) were fully achieved (14 and 15 countries, respectively, demonstrated achievement of Outcomes), Outcome 2 (sectoral and thematic policy) and Outcome 3 (capacity strengthening) were partially achieved (9 and 13 countries, respectively, demonstrated achievement of Outcomes). PAGE had excelled in facilitating knowledge exchanges through numerous events and dissemination of knowledge products. At global level, it contributed to debates about IGE as a driver for achieving sustainable development and climate goals. PAGE did contribute to increased coordination and collaboration across sectors, enabling policy coherence, and reached its objective of building capacities to strengthen IGE action.

The performance of PAGE's Operational Strategy 2016-2020 was assessed by the evaluation as 'Satisfactory'.

Sustainability

The sustainability of PAGE Outcomes during the Operational Strategy 2016-2020 period was found to have a high degree of dependency on socio-political factors, with varying success across Partner Countries in securing strong ownership and commitment on the part of the government and other stakeholders to sustain Outcomes with a generally weak mechanism put in place to adapt to changes in the social/political context.

Some factors affecting financial sustainability were within PAGE's remit, such as: the strong role of the PAGE Secretariat, room for improvement to engage Multilateral Development Banks (including potential investment plans linking downstream activities with PAGE's upstream achievements), and the timing of developing sustainability plans as part of the exit strategy from core PAGE support. The coordination of national actors, anchoring capacities in national training institutions, and the role of the Resident Coordinator Office (RCO) influenced the institutional sustainability of PAGE benefits. Stakeholders were positive about the level of country driven-ness and ownership of PAGE although they assessed other aspects of institutional sustainability more critically.

The Evaluation Team found that 83% of countries that joined PAGE in 2013 and 2014 achieved results across all four PAGE Outcomes by 2020, within a 7–8-year timeframe. However, only 40% of countries that joined PAGE in 2015 and 2016 achieved results across the four Outcomes within a 5–6-year timeframe. This evidence would suggest that a 7–8-year timeframe could be more appropriate to achieve PAGE Outcomes and is consistent with the international development context in which projects typically achieve their envisaged Outcomes beyond project close, although their Outputs are expected to indeed be delivered within the project's lifetime (through the use of those Outputs). The evaluation survey uncovered mixed stakeholder perceptions about the sufficiency of the 5-year timeframe.

Lessons learned

Unearmarked funding provides important flexibility to respond to country needs, which is particularly important for a demand-driven programme like PAGE, whose legitimacy and sustainability are anchored in national ownership.

The stability, institutional authority, competences, and personal motivation of boundary-spanning actors, like a National Coordinator, are key for building bridges between project implementers and government partners as well as for facilitating collaboration between stakeholders within national steering and governance mechanisms.

Ensuring relevance and additionality can be enhanced through deep context analysis, baseline setting, and identification of entry points and impact channels. as evidenced by PAGE's stocktaking exercise, which ideally strengthens high level political support in a context where influence on government partners is low.

Recommendations

For UNEP FMO, in collaboration with PAGE Secretariat and MPTF Office continue to channel all donor funds through the existing MPTF to enhance efficiency and effectiveness.

For PAGE Management Board and Donor Steering Committee Build high level political support in Partner Countries for IGE while also strengthening connections between PAGE's upstream normative work and follow-up action on the ground to operationalise IGE through optimising connections with donors' bilateral programmes and bringing in more influential partners (like MDBs and/or specialised UN agencies, like IFC, World Bank) strengthening these impact drivers in the programme's ToC.

PAGE Country Teams and National Coordinator, supported by Secretariat Strengthen linkages with the Resident Coordinator's Office to ensure that PAGE elements feed integrally into country programming and to leverage the Resident Coordinator's role as the UN's 'ambassador to the country' to build high level political support and enhance PAGE's external coherence.

For National Coordinators, PAGE Country Teams, and PAGE Secretariat Enforce systematic use of the traffic light system as a monitoring instrument and results-based management tool.

For PAGE Secretariat Develop and include appropriate indicators for gender mainstreaming and human rights within performance management and reporting systems. This should go beyond the work on gender mainstreaming led by partner initiatives/agencies with specialized mandate.

For PAGE Secretariat in collaboration with the Management Board, Adopt a more strategic approach; an institutionalized mandate with roles, responsibilities, appropriate support mechanisms for the involved actors, monitoring, and ongoing follow-up; using mutually understood shared criteria and a systematic process for documentation and dissemination feeding into knowledge management and to evaluate effectiveness, document good practice, to leverage peer learning and South-South exchange towards strengthening execution and sustainability.

For the UNEP Evaluation Office and Evaluation Offices of other UN agencies Considering the ambition of UN Reform: Develop relevant guidance, tools, and templates (including detailed criteria matrix) for the evaluation of interagency programmes.

Chapter 5

Compliance with Evaluation Recommendations

This chapter reports on the formal response to evaluations, and the implementation of evaluation recommendations issued. Recommendations are proposals for specific action to be taken by identified people/position-holders to resolve concrete problems affecting a project, programme or policy or the sustainability of its results. They are formulated through a rigorous approach and are distilled from evidence-based findings from evaluation processes involving project managers, evaluation managers, evaluation consultants and a wide range of stakeholders.

Evaluations aim to promote accountability and learning. The formal evaluation recommendation compliance system also reinforces the positive feedback and learning opportunities that evaluations can provide to the organization. The 'process benefits' from evaluation work can be considerable in terms of encouraging staff, partners and other stakeholders to reflect on the strengths and weaknesses of past performance, and to integrate existing good practices and ideas for improvement into future programme strategies, intervention designs and management actions.

Categorisation of recommendations

Evaluation recommendations are categorized according to their priority level, defined by the Evaluation Office in the following manner:

- **Critical** (high level recommendations): Address significant and/or widespread deficiencies in governance, risk management, or internal control processes such that likelihood of achieving the programmed objectives is greatly reduced, or the provisions of relevant principles or policies are compromised.
- **Important** (medium level recommendations): Address reportable deficiencies or weaknesses in governance, risk management, or internal control processes such that likelihood of achieving the programmed objectives is reduced, or the provisions of relevant principles or policies are threatened.
- **Opportunity for improvement** (low level recommendations): Address distinctive challenges and comprise suggestions to improve performance that do not meet the criteria of either critical or important recommendations.

Table 6 below shows the proportion of recommendations for each of these priority levels that were issued in 2022 and 2023.

Table 6. Priority level of recommendations, 2022-2023

Critical	55
Important	58
Opportunity for Improvement	43
Total	156

The Evaluation Office also characterises recommendations by type based on the 'span/locus of management' required to resolve the problem/issue, in the following manner:

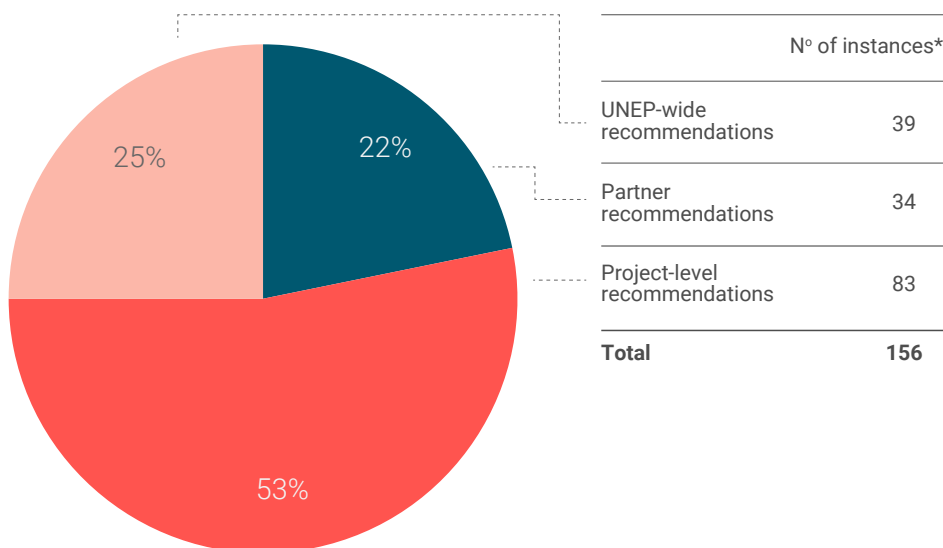
- **Project level recommendations:** Where the actions of those UNEP staff managing the evaluand, or the staff within the corresponding unit/branch can address the recommendation or the underlying problem independently.
- **UNEP-wide (institutional/cross-cutting/corporate) recommendations:** where the actions of UNEP staff/those UNEP staff managing the evaluand cannot address the recommendation or the underlying problem

independently; or where the action(s) to be taken to resolve the problem, which could have been caused by systemic issues or gaps in UNEP’s operational requirements, require approval/leadership from UNEP senior management and/or coordination among several different parts of UNEP.

- Partner recommendations: where the actions to be taken to resolve the problem require approval/leadership from UNEP partners (e.g., Executing Agencies; National Governments/Ministries; Research Organisations; Private Sector; Steering Committees; Academia; UN agencies).

Figure 33 below shows the proportion of recommendations for each of these categories. In 2022 and 2023, a total of 156 recommendations were issued comprising those addressed to projects (83) and their respective partners (34), those addressing projects but with corporate UNEP-wide implications (39).

Figure 33. Distribution of the type of recommendations issued, 2022-2023

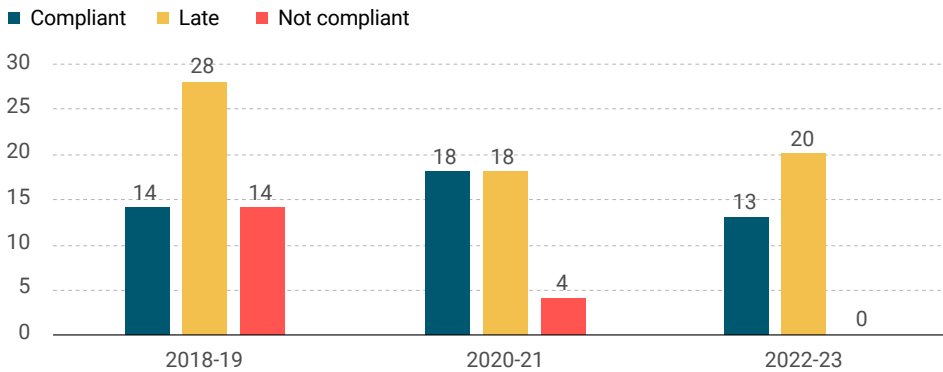


The UNEP-wide (corporate/high-level) recommendations category captures recommendations where the required management actions need to be taken by several actors and at levels beyond the management authority of the project or programme manager that received the implementation plan. The nature of such recommendations means they often take more time to implement.

As per the UNEP Evaluation Policy (2022) the findings from evaluations are to be discussed with relevant programme staff and senior management where specific evaluation recommendations will be reviewed. Subsequently, programme staff/senior management prepare a formal management response to the evaluation that specifies which recommendations were accepted, what action will be taken

and a timeline. A management response to the evaluation recommendations, in the form of a formal Recommendations Implementation Plan, should be prepared and sent back to the Evaluation Office within one month.

Figure 34. Submission of management response compliance levels, 2018-2023

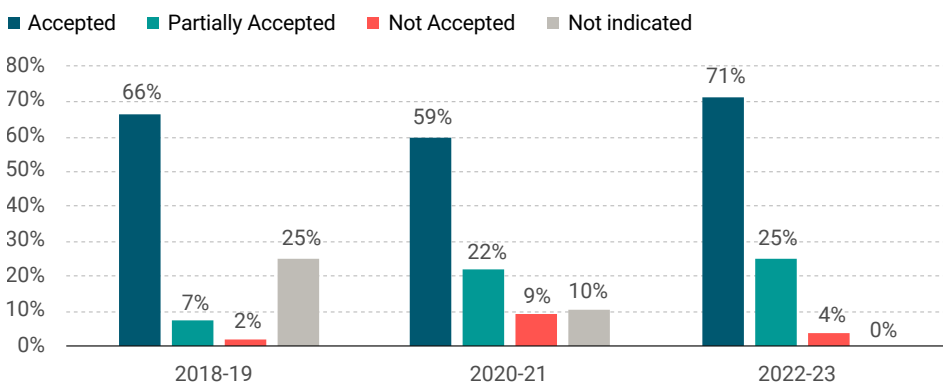


In 2022 and 2023, 34 implementation plans for evaluation recommendations were sent to UNEP substantive offices. As shown in **Figure 34** above, submission of management responses to recommendation implementation plans by the responsible project officers was at 100%. However, approximately 60% of the responses were received after the required timeframe of 1 month. Nevertheless, this is a substantial improvement on the previous two biennia.

Compliance rates for completion of management responses to evaluation should be maintained at 100% and **the timeliness of management response submission needs continued management attention.**

Below, **Figure 35** shows the overall level of acceptance by project/programme staff of the evaluation recommendations issued during the biennia of 2018-19, 2020-21 and 2022-23.

Figure 35. Overall acceptance of recommendations, 2018-2023



Partial acceptance of recommendations happens when Project Managers accept the problem/issue identified in the recommendation but propose an alternative course of action to resolve the problem/issue than that suggested by the evaluators. Reasons for non-acceptance of recommendations in this biennium have included:

- One (1) GEF project asserting that acceptance of recommendations was not possible in the current implementation context of Afghanistan ruled by the Taliban regime.

Compliance with recommendations

All recommendations must have a compliance dimension. The Evaluation Office monitors compliance of all project-level implementation plans every six months for a total period of 12 months from its finalisation. It prepares compliance assessments of the implementation of accepted recommendations, based on submitted compliance reports prepared by the organizational unit responsible for the management response. The compliance performance against the recommendations is then reported to senior management on a six-monthly basis and to Member States in the Evaluation Synthesis Report.

At each assessment point, the progress in implementing agreed recommendations is assessed by the Evaluation Office. The compliance level is based on implementation progress updates and supporting evidence provided by the responsible programme staff.

Recommendations are deemed to be:

- **Compliant** (fully implemented)
- **Partially compliant** (only partially implemented)
- **Not compliant** (not implemented)
- **No further action required** (if events overtake what is planned/required)

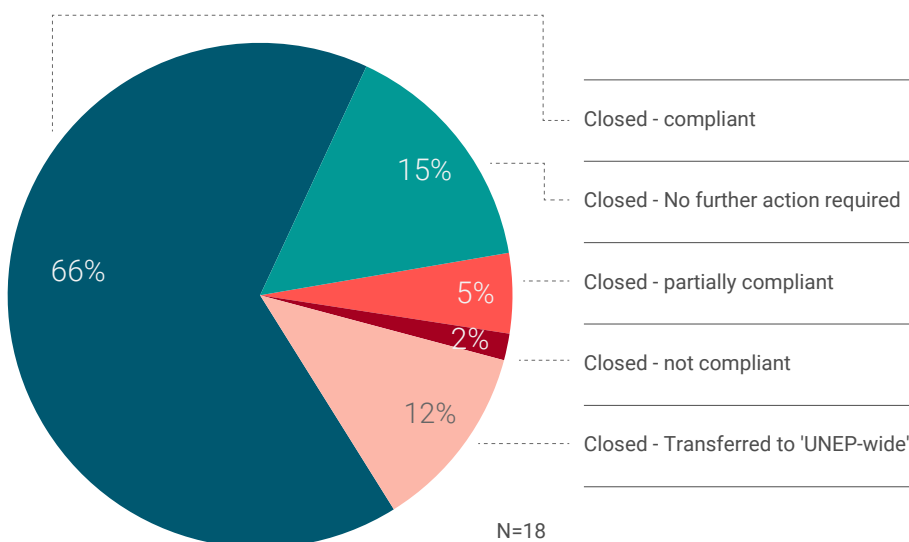
UNEP cannot enforce compliance for any recommendation made to a third party unless there is a contractual/legal agreement in force. If there is no binding legal agreement with the third party in place, UNEP can only communicate the recommendation to the third party for their consideration.

When 12 months from the date of finalisation of the recommendation implementation plan have elapsed, the implementation status of all recommendations is updated and the compliance status recorded. All recommendations in the plan are then recorded as 'closed'. This means that staff responsible for implementing the evaluation recommendation have

12 months from the date the implementation plan is received by the Evaluation Office to implement them fully.

Figure 36 below shows the compliance rates across UNEP Divisions in the cases where evaluation recommendation implementation plans reached their time limit or were fully implemented in 2022 and 2023.

Figure 36. Compliance status of recommendation implementation plans closed, 2022-2023

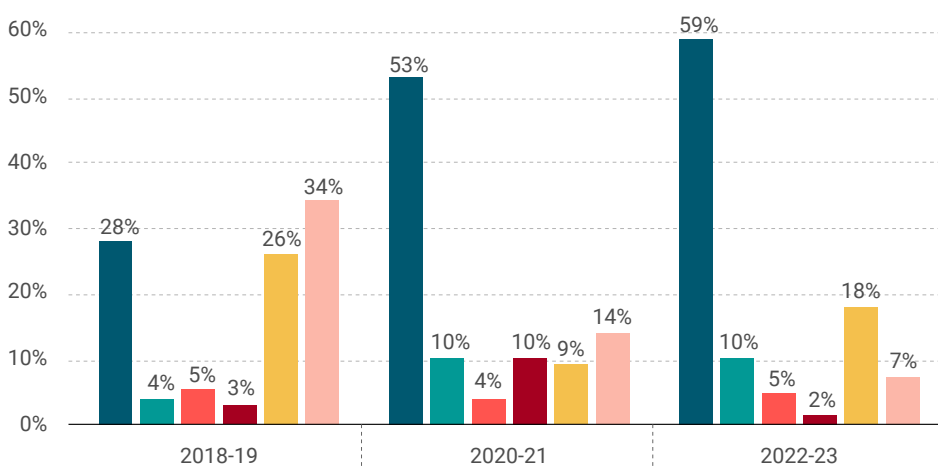


In **Figure 37**, 'open recommendations' are those that have yet to be fully implemented and have yet to reach their 12-month implementation deadline. The chart reflects an improvement in evaluation recommendation compliance this biennium with the proportion of recommendations closed as 'Compliant' or 'No further action required' increasing to 69% from 63% in the last biennium. In addition, the proportion of recommendations closed as 'not compliant' fell from 10% in 2020-21 to 2% in 2022-23. The Evaluation Office attributes much of the improvement to the use of corporate dashboards that are reviewed by the Senior Management Team every quarter.

The Programme of Work 2022-2023 performance indicator for recommendation compliance is the *"Percentage of accepted evaluation recommendations implemented within the time frame defined in the implementation plan"*. The target level expected by December 2023 was 85%. The actual compliance level achieved was 92%²⁹. **Recommendation compliance levels need continued management attention to meet UNEP's PoW performance targets.**

²⁹ The percentage includes the category 'No further action required' and excludes a category "Closed - Transferred to UNEP-Wide"

Figure 37. Comparison of recommendation compliance status, 2018-2023



Overall compliance status

■ Closed - Compliant	74	120	6
■ Closed - No further Action Required	11	23	2
■ Closed - Partially Compliant	13	9	23
■ Closed - Not Compliant	88	33	13
■ Open - Partially Compliant	67	21	75
■ Open - Not Compliant	88	33	13
Total ³⁰	261	228	128

³⁰ The data excludes a category termed as "Closed - Transferred to UNEP-Wide"

Chapter 6

Special Study – A Synthesis of findings from Assessments of Project Design Quality

This chapter explores trends and patterns in information collated by the Evaluation Office on the quality of UNEP project designs. It also examines links between assessments of project design and the later performance of the same projects as assessed through evaluation. The work was undertaken in 2022 based on data collated between 2017 and 2021.

Background

UNEP has two separate mechanisms that assess the quality of project designs; one under the auspices of the Project Review Committee (PRC) prior to project implementation at the point of project approval, and another that forms a part of project level performance assessments (evaluation/review) at project completion. The PRC Secretariat is hosted within Programme Coherence/Assurance Unit (PCAU) of UNEP's Policy and Programme Division (PPD) and serves the PRC which is mandated to review and assess the quality of projects before their consideration for approval³¹. The PRC assessment of the quality a new Project Document (ProDoc) has a specific focus on the project's logic and approach to achieving results. Technical reviewers assess each ProDoc prior to a PRC meeting and make use of the quality of project design matrix and checklist.

³¹ It is noted that the PPD presented its own study on the quality of project designs in early 2022.

The PRC Secretariat has used the project design matrix tool to assess the quality of project designs since 2013. The assessment approach used by the PRC was originally developed by the Evaluation Office but has been modified over the intervening years. Nevertheless, many similarities between the PRC assessment tool and that used by the Evaluation Office still remain.

As a part of UNEP's standard project evaluation/review process, the quality of a project design of the most recently formally approved/revised project document is assessed using a standard tool. During the evaluation inception phase, ratings are attributed to a standard set of criteria, using a six-point scale (**See Table 2**) and an overall Project Design Quality (PDQ) rating is established. The completed PDQ assessment forms a part of the Inception Report of every evaluation and the overall PDQ rating is included in the final evaluation report along with a summary of the project's design strengths and weaknesses.

In 2019, the Evaluation Office undertook a small study to understand whether the Quality of Project Design ratings had any correlation with evaluation ratings for the Effectiveness criterion (both at an overall level and also at the levels of the Effectiveness sub-criteria; Availability of Outputs, Achievement of Outcomes and the Likelihood of impact). The hypothesis was that a more favourable assessment of the Quality of the Project Design might be associated with higher levels of performance under the Effectiveness criterion.

Based on a sample of 49 project evaluations carried out in 2018 and 2019, the study found³², that there was no statistically significant correlation between evaluation ratings for the Quality of Project Design and any aspect of Effectiveness. This suggests that there are numerous factors influencing the effectiveness of project interventions and that the quality of project design is not, on its own, a sufficiently strong factor to directly influence the level of effectiveness of a project.

Although no statistically significant correlation was found, frequency patterns showed that the predictive power of quality of project design ratings was stronger in relation to the sub-criterion Availability of Outputs and gradually weakens in relation to the sub-criteria Achievement of project Outcomes and Likelihood of impact. The study looked at exceptional cases, particularly those where the rating for the quality of project design was low but ratings for effectiveness were relatively high and vice versa. It found that political will and country ownership were critical factors in explaining situations where quality of project design and project effectiveness assessments diverged.

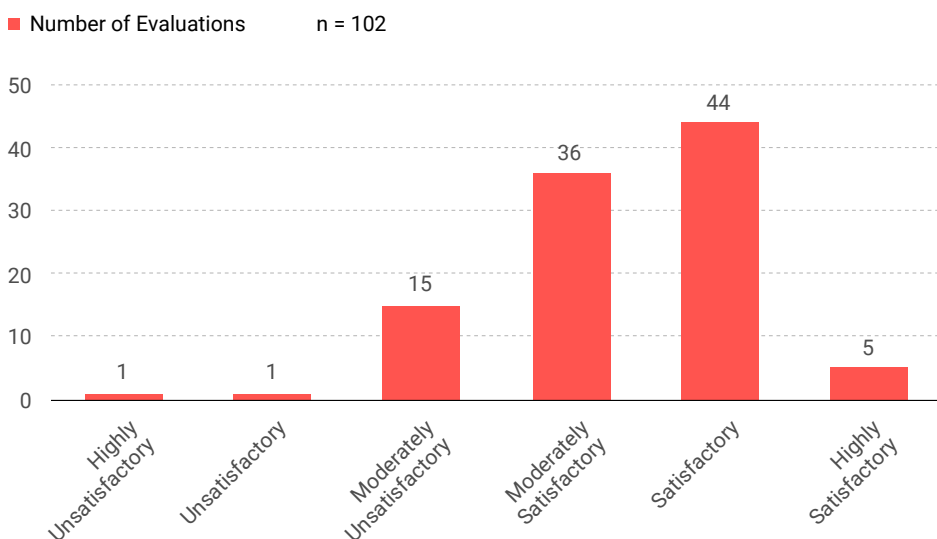
³² This work is summarized in the 2018-19 Biennial Evaluation Synthesis Report, page 24

In 2022, as a follow on from this earlier work, the Evaluation Office examined data from the assessment of project design quality (PDQ) presented in 86 evaluation inception reports in combination with the respective data on project performance collated from the corresponding final evaluation reports completed between 2017–2021.

Analysis of project design quality assessments in evaluated projects

FINDING 1: In 49% of cases the PDQ assessment determined that the project design was rated as S/HS and 81% of projects were assessed as being in the ‘Satisfactory’ range (i.e. either HS, S or MS). This suggests that the quality control/ approval process at project design stage is generally effective and the potential for improving project designs must be identified at a more detailed level. **Figure 38** below shows that out of 105 evaluations, 17 projects were in the ‘Unsatisfactory’ range, 85 in the ‘Satisfactory’ range and 3 with no ratings.’

Figure 38. Range of ratings for project design quality assessments, 2017-2021



FINDING 2: Given that the assessment tools have been constant since 2018, over time there appears to be a general tendency for the quality of project designs to have improved: PDQ assessments from 2018 and 2019 are concentrated between ‘Moderately Unsatisfactory’ and ‘Satisfactory’, the 2020 PDQ assessments show a wider spread and the 2021 PDQ assessments are concentrated between ‘Moderately Satisfactory’ and ‘Highly Satisfactory’ (i.e. fully in the ‘Satisfactory’ range).

Figure 39. Range of PDQ ratings, 2018-2022

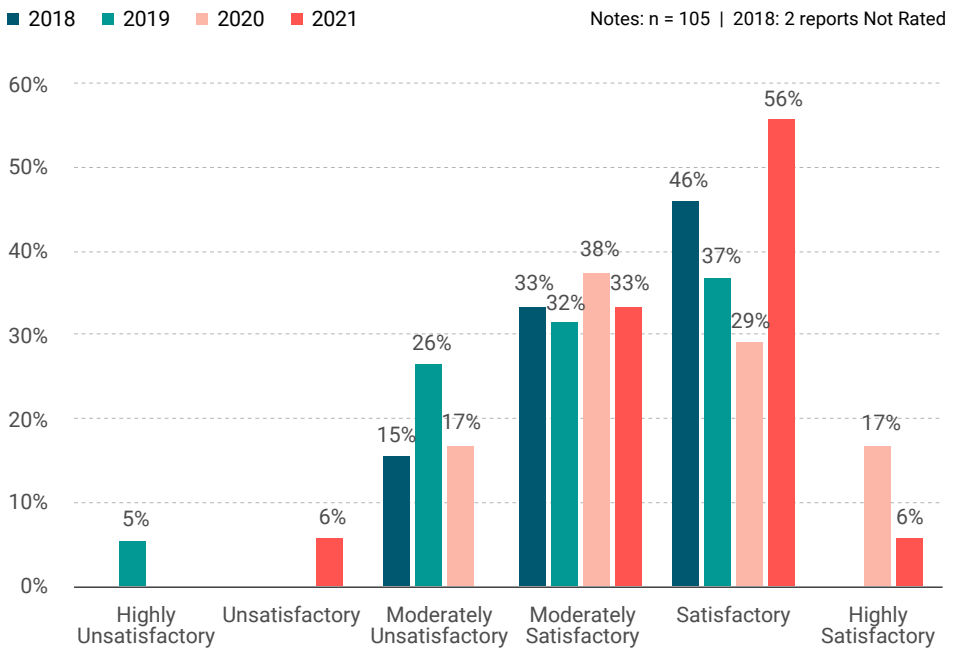
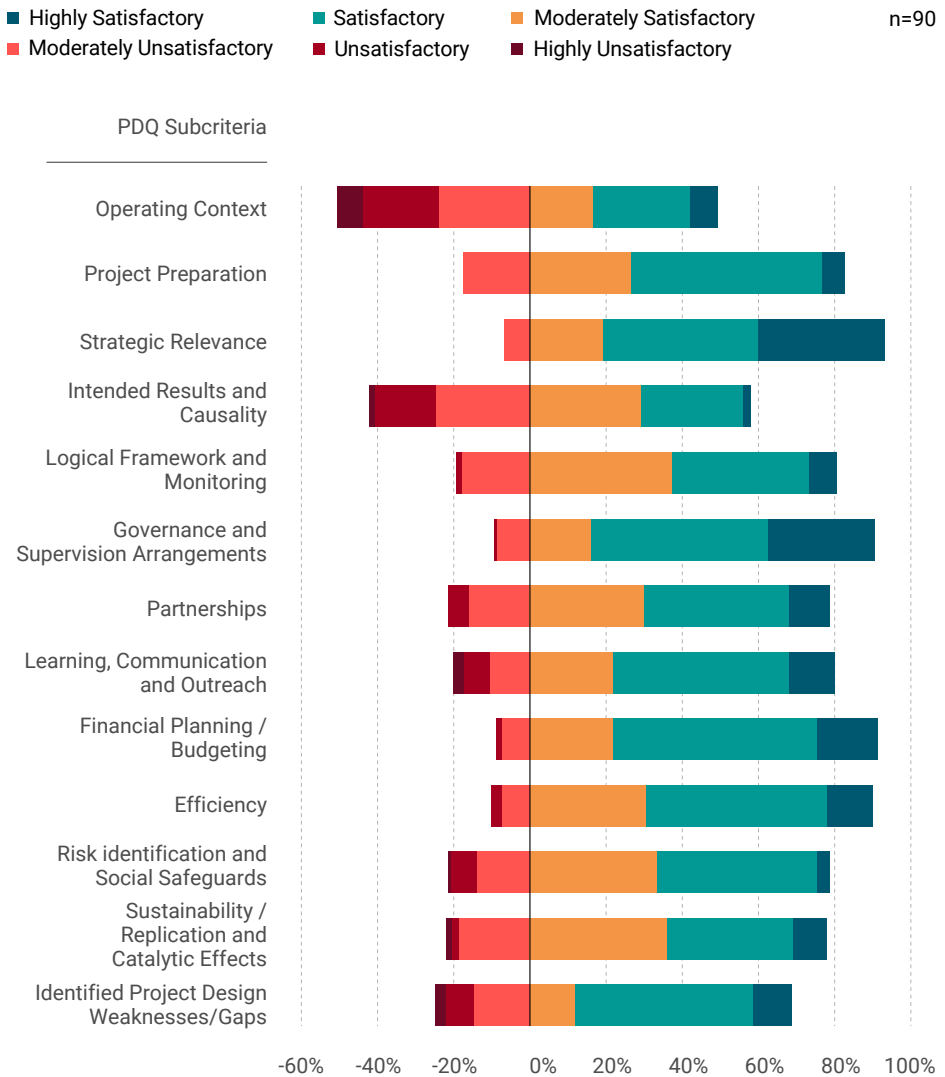


Figure 39 above shows the overall distribution of PDQ ratings by year. The 39 projects evaluated in 2018 had PDQ ratings concentrated between ‘Moderately Unsatisfactory’ to ‘Satisfactory’; the 19 projects evaluated in 2019 were rated between ‘Highly Unsatisfactory’ and ‘Satisfactory’ while those (24) in 2020 were rated between ‘Moderately Unsatisfactory’ to ‘Highly Satisfactory’ and the 18 projects evaluated in 2021 had ratings ranging from ‘Unsatisfactory’ to ‘Highly Satisfactory’. Note, the year of design approval for these projects would be approximately 3–5 years prior.

FINDING 3: The variations within project design aspects show that dimensions with weaker design ratings include: ‘Intended Results and Causality’; ‘Sustainability/Replication and Catalytic Effects’; ‘Risk Identification and Social Safeguards’ and ‘Learning, Communication and Outreach’.

It can be seen from **Figure 40** that aspects of ‘Intended Results and Causality’ as well as ‘Identified project design weakness and gaps’ have the highest proportion of projects with HU/U ratings, followed by Sustainability, Replication and Catalytic Effects, Learning, Communication and Outreach and Risk Identification and Social Safeguards sections of the PDQ assessment. The black vertical line emphasising the boundary between the ‘Satisfactory’ and ‘Unsatisfactory’ ranges.

Figure 40. Comparison of project design quality sub-criteria ratings



The 'Intended Results and Causality' assessment covers whether the project had well-defined causal pathways from project Outputs to Outcomes and through to impacts either in its logical framework or Theory of Change, including impact drivers, assumptions, and roles of key stakeholders for each causal pathway at project design. It also includes the assessment of whether the result statements are realistic with respect to timeframe and scale of the intervention. Projects in the 'Unsatisfactory' range tended to have; poorly formulated result statements where activities, Outputs and Outcomes were often used interchangeably, lacked clear descriptions of causal pathways, and assumptions, drivers or roles of key stakeholders were weakly addressed.

These projects also had unrealistic timeframes especially those that were designed for legal frameworks to be adopted at national levels.

Assessment of 'Sustainability, Replication and Catalytic Effect' covered whether a project design addressed socio-political, financial, institutional and environmental sustainability issues; whether there was a credible strategy to sustain project outcomes/effects or an appropriate exit strategy; and whether the project design presented strategies to promote or support scaling up, replication and/or catalytic action. Projects in the 'Unsatisfactory' range lacked a credible strategy to help sustain a project's effects, had no exit strategy and often relied solely on funding from donors to continue the work.

Assessment of 'Learning, Communication and Outreach' covered whether the project design had a clear and adequate knowledge management approach, whether appropriate channels of communication with key stakeholders had been identified for the duration of the project and what plans were in place for dissemination of results and lesson sharing at the end of the project. Projects with HU/U ratings did not have an adequate knowledge management approach or have plans in place for dissemination of results and lessons sharing at the end of the project or lacked all three factors at project design.

Assessment of 'Risk Identification and Social Safeguards' covered whether risks were properly identified as part of the Theory of Change/Logical Framework and in a risk table, whether potentially negative environmental, economic and social impacts of the project had been identified and the associated mitigation strategy adequate, as well as whether the project had adequate mechanisms in place to reduce its negative environmental footprint. Projects with HU/U ratings either had no risk analysis as part of the ToC/Logical framework nor in a risk table or lacked a combination of having no risk analysis table or analysis of potential negative environmental, economic and social impacts of the project at the design stage.

FINDING 4: It appears that the role of the PRC in providing inputs to improve the quality of project design is not being fully utilized (identified project design weaknesses/gaps). This may be either because design weaknesses are not being identified by the PRC or that the advice provided by PRC assessments is not being adopted by Project Managers.

The assessment of Identified 'Project Design Weakness and Gaps' covered two aspects: whether issues raised by the PRC were addressed or not, and whether there were any major issues not flagged by the PRC. On further analysis of the PDQ assessment by the Evaluation Office, it was found that at least nine projects had a rating of U/HU; The main reason for the HU rating was that there was lack of evidence that any action had been taken by the project team to address the

issues raised by the PRC review. For the projects rated as 'Unsatisfactory', one lacked evidence of a PRC review, one was due to major issues not raised by PRC and the remaining five were a combination of issues not being addressed as well as issues not raised by the PRC. Seven projects had 'No rating' for this section due to no evidence being presented at the evaluation inception phase.

FINDING 5: As the assessments of 'Logical Framework and Monitoring' show higher design quality ratings than those for 'Intended Results and Causality', this suggests that although the needed elements of the project are being considered, they are not sufficiently embedded in a clear and comprehensive articulation of intended change processes.

Figure 41. Comparison of project design quality sub-criteria with 'Satisfactory' rating as the minimum acceptable standard

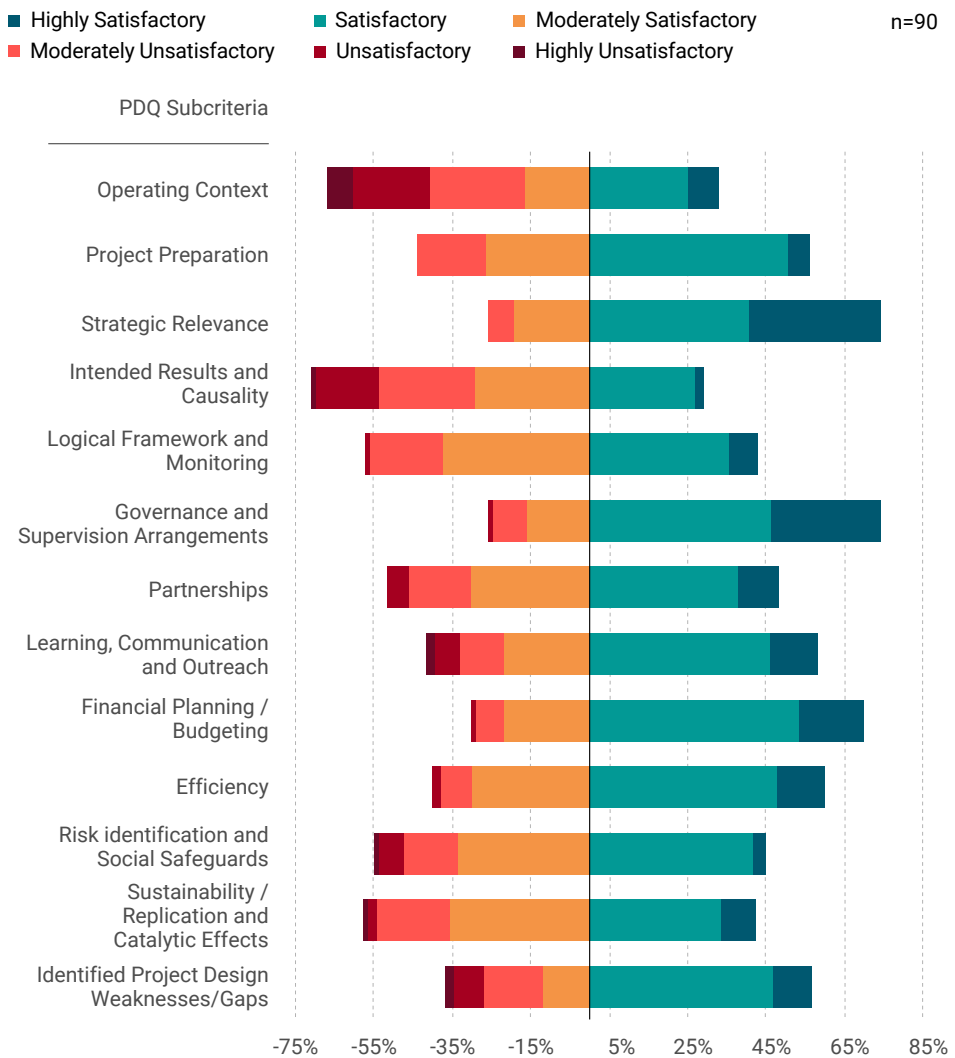


Figure 41 shows data based on the assumption that the ‘Satisfactory’ level would be the minimum PDQ standard acceptable for a project to be given UNEP approval (as emphasised by the vertical black line in Figure 41. ‘Intended Results and Causality’ and ‘Sustainability, Replication and Catalytic Effects’ have the highest proportion of ratings below the ‘Satisfactory’ level, followed closely by ‘Risk Identification and Social Safeguards’ and ‘Logical Framework and Monitoring’. This highlights the dimensions where the potential ‘payoff’ is greatest for increased attention in project design at approval.

Comparing project design quality ratings with evaluation project performance ratings

An additional comparison was made between the ratings from the PDQ and the performance ratings attained during the Terminal Evaluation of the same project. This was analysed for the criteria which have the highest weightings in the calculation of overall project PDQ scores, i.e. Project Preparation, Intended Results and Causality, and Sustainability / Replication and Catalytic Effects. It is noted that this is in no way a pre and post project implementation assessment and the findings are both tentative and of a qualitative nature and should be interpreted with caution.



Project preparation (PDQ) vs preparation and readiness (TE)

Higher rating at TE:

Four projects had ratings that were in the ‘Unsatisfactory’ range in the PDQ assessment but in the ‘Satisfactory’ range at evaluation. Reasons for performance ratings in Terminal Evaluation assessments having a better rating than the corresponding criterion in the PDQ assessment included: the project document not presenting a clear and adequate problem analysis; poor situation analysis; inadequate stakeholder analysis/engagement; and Human Rights, Gender or Indigenous People’s rights perspectives not being covered at the design stage.

Lower rating at TE:

The sample included 17 projects where ratings changed from the ‘Satisfactory’ range in the PDQ assessment to the ‘Unsatisfactory’ range at evaluation. Reasons for Terminal Evaluation assessments having a lower rating than comparative PDQ assessment included:

- Unrealistic results frameworks;
- Legal agreements not in place at project inception phase;
- Project inception phase not being utilised properly (timely hiring of project team, formation of steering committee, key stakeholder engagement);

- Inaccurate stakeholder mapping;
- Long time (more than 1 year) elapsed between project approval and onset of project implementation;
- No adaptive management (due to changes in statutory requirements, funding and budget revisions, inaccurate baseline conditions, etc).

The above factors usually occurred in combination to varying degrees rather than in isolation.



Intended results and causality (PDQ) vs effectiveness (TE)

Higher rating at TE:

The sample included 24 projects with ratings changed from the 'Unsatisfactory' range in the PDQ assessment to the 'Satisfactory' range at evaluation. Reasons for performance ratings in Terminal Evaluation assessments having a higher rating than the corresponding criterion in the PDQ assessment included:

- The project document having no Theory of Change or weak logical frameworks, no or weak descriptions of causal pathways, absence of drivers and/or assumptions being identified in TOCs as well as weak connection between stakeholder mapping and causal pathways;
- Reconstruction of ToC at Evaluation clarifying the results which meant the evaluation assessment was able to pick up the Outputs and Outcomes that were delivered;
- Local context – nature, behaviour, culture and customs of local stakeholders were not evident in the design but were being taken into account during implementation;
- Knowledge products produced in local and official languages of the country;
- Engagement with national partners, private sector, marginalised and disadvantaged people;
- Strong political will/buy in (from multiple Ministries);
- Commitment of executing agency/partners and good supervision;
- Lasting social conflict resolution ;
- Increased local, national and global awareness of project aims;
- Regulatory and institutional frameworks in place;
- Adaptive management;
- Complementarity and synergies with ongoing projects rather than duplication.

Lower rating at TE:

The sample included four projects with ratings that changed from the 'Satisfactory' range in the PDQ assessment to the 'Unsatisfactory' range in the corresponding Terminal Evaluation. Reasons for performance ratings in terminal evaluation assessments having a lower rating than the corresponding criterion in the PDQ assessment included:

- Local context – nature, behaviour, culture and customs of local stakeholders not taken into account;
- No adaptive management and/or lack of documentation for changes made;
- Key stakeholders (especially private sector) not engaged;
- Lack of political will;
- Delays in starting the project.



Sustainability / replication and catalytical effects (PDQ) vs sustainability (TE)

At terminal evaluation, the overall sustainability rating is taken as the lowest rating of the three sub-criteria; socio-political, financial and institutional sustainability.

Higher rating at TE:

There were three projects where ratings changed from the 'Unsatisfactory' range in the PDQ assessment to the 'Satisfactory' range in the Terminal Evaluation. The reasons for terminal evaluation assessments having a higher rating than the comparative PDQ assessment included:

- An exit strategy was formulated during the project implementation;
- In at least one project, the socio-political sustainability was associated with the higher rating at evaluation compared to design. This project had generated strong ownership, interest and commitment among government and other stakeholders.

Lower rating at TE:

There were 16 projects where ratings moved from the 'Satisfactory' range in the PDQ assessment to the 'Unsatisfactory' range at the time of Terminal Evaluation. Reasons for terminal evaluation assessments having a lower rating than the respective PDQ assessment included:

- Financial sustainability being least likely – either further funding was needed to sustain results, or the funding mechanism described in the project document was not realised, including co-finance, or the project did

not address the shortfall in funding being further compounded by absence of exit strategies;

- Institutional sustainability was another weak link. The projects were reliant on institutional frameworks set up specifically for the project and with no exit strategy for handover, led to a lower rating;
- None of the projects assessed had an exit strategy at project design nor was one developed during project implementation.

Overall, this study on the quality of project design provides some valuable feedback to inform UNEP staff of areas for potential improvements to project design, project approval processes and project implementation.

Chapter 7

UNEP's Updated Evaluation Policy and Strategy

The context in which UNEP operates has changed significantly in recent years, dominated by three interconnected crises—climate change, biodiversity loss and pollution. All three are driven by unsustainable production and consumption patterns, which put global economic and social well-being at risk and undermine the environmental foundations that provide opportunities to reduce poverty and improve lives. This already challenging global situation was further complicated by the rise of nationalism threatening multilateralism, armed conflicts and the COVID-19 pandemic causing disruptions to global markets, international political processes, peace and security.

UNEP articulated its intended response to this situation with the development of its Medium-Term Strategy 2022-2025. The MTS aims to drive “transformative, multi-stakeholder actions that target the causes of climate change, biodiversity loss and pollution. In doing so, UNEP aims to ensure that science remains at the centre of decision-making processes, and that environmental rule of law continues to underpin global environmental governance. UNEP also aims to boost efforts to use digital technologies for inclusive, transparent and innovative outcomes.”

In addition, in recent years, the UN Secretary-General has overseen an extensive UN reform effort and in his report A/72/492 on “Shifting the management paradigm in the United Nations: ensuring a better future for all”, indicated his intention to strengthen the evaluation capacity within the UN Secretariat to better inform programme planning and reporting on programme performance. He stressed that results of evaluations are to be used by programme managers to better plan and adjust their activities. As such, evaluation is integral to learning and supports improved results-based management and increased transparency on programme delivery to Member States.

The role of the evaluation policy

The first UNEP Evaluation Policy was approved in 2009, updated in 2016 and fully revised in 2022 in-line with current good practice, new UN Secretariat requirements and to reflect the new strategic context in which it sits. The Evaluation Policy 2022 recalls the mandates for evaluation, sets out the purpose of evaluation and describes the organization and management of the evaluation function, including what evaluation types are carried out, what products are generated, the principles that guide evaluative work and which evaluation criteria are assessed. The Evaluation Policy also outlines the general processes by which it is operationalized, how evaluation findings and recommendations are to be used and how the performance of the evaluation function itself will be assessed.

The evaluation approach in UNEP is driven by the policy, but also pays due regard to other UNEP policies and strategies, including Gender, Human Rights, Monitoring, Partnership, Safeguards, Disability and Inclusion and the Bali Strategic Plan for Technology and Capacity Building. The operationalization of the Evaluation Policy is further specified in the Evaluation Manual which outlines UNEP’s performance assessment processes and in the UNEP Programme and Project Management Manual, which covers the roles of many divisions and offices involved in the project cycle and the implementation of the MTS.

Five strategic and interrelated priority areas

The Evaluation Office has introduced a comprehensive strategic approach to assess UNEP’s performance and enhance organizational learning, evidence-based decision-making, and transparency. The strategy encompasses five interrelated priority areas: Developing and maintaining a strategic evaluation agenda, enhancing use of evaluation throughout the house, maintaining a focus on quality of evaluation processes and products, building capacity across the house for robust project-level performance assessments (management-led review), establishing a stable resource base and funding approach to support the evaluation function.

The priority to develop and maintain a strategic evaluation agenda is a commitment to:

- Aim for balanced coverage of thematic areas in alignment with UNEP's Medium-Term Strategy.
- Increase focus on strategic evaluations, allocating more staff time and financial resources.
- Selectively evaluate impact, introducing Impact Evaluations or Studies at a minimum rate of 1 per biennium.
- Purposively sample completing projects for evaluation coverage, emphasizing quality over quantity.

The second priority, enhance use of evaluation throughout the house, is a commitment to:

- Promote acceptance and implementation of evaluation recommendations.
- Facilitate interactive dialogues between evaluators and managers for better ownership and utility of findings.
- Introduce UNEP-wide recommendations, fostering management dialogues and recording narratives.
- Communicate evaluation findings through various channels, emphasizing strategic evaluations.

The third priority is to maintain a focus on quality of evaluation processes and products, and is a commitment to:

- Establish and maintain robust evaluation quality assurance processes.
- Designate an Evaluation Manager and internal Peer Reviewer for each evaluation.
- Apply quality standards at key evaluation stages, with a focus on clear identification and feasibility of recommendations.

The priority to build capacity across UNEP for robust project-level performance assessments (management-led review) is a commitment to:

- Provide systematic support for Project Managers conducting Terminal Reviews.

- Prepare training materials, establish Project Performance Assessment Focal Points, and conduct training webinars.
- Monitor the quality of Management-led Terminal Reviews over time.

The final priority is to establish a stable resource base and funding approach to support the evaluation function, and addresses the need to:

- Ensure adequate resourcing for the efficient operation of the evaluation function.
- Introduce a funding principle that costs of evaluation for projects and programmes supported by extrabudgetary resources should be covered by those resources.
- Set a funding mechanism and reflect it in the UNEP Programme and Project Management Manual.

Key Performance Indicators (KPIs) are outlined for each priority, tracking progress and effectiveness in achieving the strategic priorities. The strategy aims to strengthen evaluation processes, enhance communication, and maintain high-quality standards across UNEP's activities.

Chapter 8

Performance of the Evaluation Office and Evaluation Quality

The OIOS evaluation dashboard study of the UNEP Evaluation Office

In May 2023, the Office of Internal Oversight Services of the United Nations Secretariat (OIOS) conducted an assessment of evaluation capacity and practice for every entity in the UN Secretariat, this was published through their ‘Evaluation Dashboard’ report. The report noted that *“UNEP has all the elements of a strong evaluation system in place”*.

The report identified areas where the Evaluation Office has made improvements since the last assessment, which included:

- Evaluation planning
- Quality of reports, 81% ‘Good’ or ‘Very Good’ (up from 64%)
- Quality of recommendations, 76% of reports with ‘Good’ or ‘Very Good’ recommendations (up from 45%)
- Reports that meet UN-SWAP requirements 23% (up from 18%), **however this is still an area requiring improvement.**

The report highlighted that the UNEP Evaluation Office has produced the highest number of evaluation reports (41) in the UN Secretariat in 2020 and 2021. The second most productive function produced 26 reports.

In addition, OIOS prepared the report A/78/70 to the General Assembly entitled “Strengthening the role of evaluation and the application of evaluation findings on programme design, delivery and policy directives”. The report found that UNEP has prepared evaluation reports of excellent quality showing clear linkages between findings, conclusions and recommendations and including cross-cutting issues, (e.g. a thorough analysis of gender considerations, human rights issues and the environment).

UN system-wide Action Plan on gender equality and the empowerment of women (UN-SWAP)

The UN-SWAP (UN system-wide Action Plan) is a UN system-wide accountability framework designed to measure, monitor, and drive progress towards a common set of standards to aspire and adhere to for achieving Gender Equality and Empowerment of Women (GEEW). The performance against these standards is assessed through 15 GEEW performance indicators for tracking six main elements of gender mainstreaming: accountability, results-based management, oversight, human and financial resources, capacity, and knowledge exchange and networking.

The UN-SWAP method for measuring performance against the indicators was revised with a more specific sub-criteria by which to examine the evaluations. A new set of indicators was applied from 2018, referred to as UN system-wide Action Plan 2.0, which needs to be considered when comparing the results prior to and since, 2018. The oversight element of the UN-SWAP 2.0 includes 3 performance indicators, including one “Evaluation Performance Indicator (EPI)” that is linked to meeting the gender-related UN Evaluation Group (UNEG) Norms and Standards and demonstrating effective use of the UNEG guidance on integrating gender in evaluation.

The EPI also calls on all reporting UN system entities to conduct at least one evaluation to assess corporate performance on gender mainstreaming every 5-8 years. This might constitute, but not be limited to, corporate evaluation of gender policy, mainstreaming, and strategy or equivalent. On the latter, between August 2021 and January 2022, the Evaluation Office conducted an “*Independent Evaluation of UNEP’s Policy & Strategy for Gender Equality and the Environment 2015-2020*”. The evaluation report was distributed in June 2022 with 12 recommendations, including on the use of UN-SWAP as an accountability and management instrument.

The UN-SWAP requires all UN entities to self-assess and report on their plan of implementation. Fourteen out of the 15 indicators assess the progress made within the institution in mainstreaming gender-responsive strategies and systems, the evaluand being the institution itself, including its systems and the actions it has taken to promote gender equality and the empowerment of women. The scores are based on a self-assessment guided by a common approach developed by UN Women in conjunction with the UN Evaluation Group. The one exception within the set of 15 indicators is the Evaluation Performance Indicator (EPI), which is assessed by an external consultant who reviews the coverage of GEEW in all the evaluation reports completed in the preceding year. A five-point rating scale (not applicable; missing; approaches requirements; meets requirements and exceeds requirements) is common to all the fifteen UN-SWAP Performance Indicators, including the EPI. A score is applied overall for the UNEP's evaluation function.

Concerning the review of project evaluations, in 2022, the UNEP Evaluation Office engaged an external consultant to undertake a review of the 16 project evaluation reports completed this year. In 2023, 28 project evaluation reports were assessed and the review outcome will be submitted to UN Women in January 2024.

In 2022, UNEP was assessed at the 'Meet Requirements' level, with a slight improvement in the score year on year. Of the 16 evaluations reviewed, 12 'Meet Requirements', 4 'Approach Requirements' and none 'Misses Requirements'. This indicates that 75% of all evaluations reviewed met the requirements in accounting for and addressing GEWE considerations; this is higher than in 2021, when only 55% of all evaluations met requirements.

The following strengths and weaknesses in evaluation reports were identified from a broader perspective.

The **main achievements** in the 2022 reports included the following:

- All the evaluations examined gender responsiveness.
- All but one evaluation provided gender-disaggregated data.
- Most evaluations included gender-relevant recommendations and/or lessons learned.
- There was more contextual gender information provided than in the past.
- Ethical standards and statements on integrity and respect for confidentiality were provided in all but two evaluations.
- Technical and institutional projects are making better links between the project and interactions with women.

- Evaluations are increasingly referring to gender in the evaluation methodology.
- There is improvement in terms of evaluators identifying the gaps of whose voices they have not been able to include sufficiently.

The **main challenges** that appeared in the 2022 evaluation reports include:

- Difficulty accessing marginalized voices and triangulating them clearly in the report, particularly given travel restrictions.
- Women can be represented as a monolith; gender-disaggregated data does not always account for socioeconomic/power differentials and ethnic/ability/livelihood or other differences.
- Unintended or unanticipated impacts on women are generally lacking.
- Intersectional analyses and/or background are still weak despite improvement in providing more contextual information. In the case of technical projects, there is a need to contextualize how the subject issue at hand may intersect with gender, class, ethnicity, or other social groupings. For instance, if the evaluation is focused on an international framework or policy, the evaluator may want to tease out how framework/policy outcomes could have a differentiated impact on women, based on the intersection of their identities. For projects that have fieldwork, evaluators may require guidance on the extent of information and type of analysis they should provide to meet these criteria.

In 2023, UNEP was assessed at the 'Approaches Requirements' level. This score is lower than the previous year. There were significantly more evaluations in 2023, a total of 28 compared to 16 in 2022, which is a 43% increase. Of the 28 evaluations, 12 evaluations were scored as 'Meet Requirements', 14 'Approach Requirements' and 2 'Miss Requirements'. This means that overall, 43% of evaluations met requirements in accounting for and addressing GEEW considerations. This is far lower than 2022 when 75% of all evaluations met requirements.

Key achievements in the 2023 evaluation reports included the following:

- 61% of evaluations met the gender related UNEG Norms and Standards and demonstrating effective use of the UNEG Guidance on integrating human rights and gender equality during phases of the evaluation.
- There were pockets of excellence, with four evaluations scoring 9 and four scoring 8. These evaluations provided a comprehensive, transformative, and intersectional gender analysis; raising gender implications

for users, beneficiaries, or women or vulnerable groups potentially negatively affected by the project.

- Gender was generally referred to in the methodology, most often referencing the UNEP Terms of Reference, and providing gender disaggregated lists of key informants.
- Most evaluations mentioned the ethics statement in the Terms of Reference.
- Most respondent data was gender disaggregated, and the ratio of women in the sample was noted in the text of most evaluations, along with respondents' institutional affiliation.

The main challenges that appeared in the 2023 evaluation reports included:

- Although all evaluations did provide an assessment against the requirement in the Terms of Reference for responsiveness to gender and human rights, in general, the evaluations were less gender-responsive than in 2022, where all evaluations displayed good gender awareness.
- Most evaluations lack discussion on approaches to intersectionality, a clear methodology, or clarification on how gender implications or improvements or changes in women's lives would be measured.
- Unintended effects on women were inadequately captured in most evaluations. However, the excellent evaluations did investigate in detail the different implications for women, vulnerabilities around gender, and the ways in which programmes had optimized benefit for and participation of women.
- Twelve of the 28 evaluations did not make any recommendation regarding gender. Of the 16 where there was at least one recommendation, 11 were deemed to be thoughtful, feasible, and gender-transformative, while the other 5 seemed to be largely compliance-based.
- There was very little to report on in the evaluations on UNEP's incorporation of the needs and interests of those living with disabilities into project-level evaluations.

Annex 1

List of evaluations in the Evaluation Synthesis Report 2022-2023

2022 Evaluation Reports

Terminal Evaluation of the UNEP/GEF project “Support to Eligible Parties to Produce the Sixth National Report to the Convention on Biological Diversity, CBD 6NR”, January 2022. Covering:

1. Support to Eligible Parties to Produce the Sixth National Report to the CBD – (Global: Africa-3, Maldives, Nicaragua, Pakistan and Solomon Islands); GEF 9832.
2. Support to Eligible Parties to Produce the Sixth National Report to the CBD (Africa-1); GEF 9817.
3. Support to Eligible Parties to Produce the Sixth National Report to the CBD (Africa-2); GEF 9824.
4. Support to Eligible Parties to Produce the Sixth National Report to the CBD (Europe, CIS and Mongolia); GEF 9822.
5. Support to Eligible Parties to Produce the Sixth National Report to the CBD (Pacific) GEF 9823.

6. Terminal Evaluation of the UNEP/GEF project “Enabling South Africa to Prepare its Third National Communication (TNC) and Biennial Update Report (BUR3) to the UNFCCC” (GEF 5237), February 2022.
7. Terminal Evaluation of the UNEP/GEF project “Mainstreaming Agrobiodiversity Conservation and Use in Sri Lankan Agro-Ecosystems for Livelihoods and Adaptation to Climate Change” (GEF 4150), May 2022.
8. Terminal Evaluation of the UNEP/FAO/GEF project “Mainstreaming Biodiversity Conservation and Sustainable Use for Improved Human Nutrition and Well-Being, (GEF 3808), April 2022.
9. Terminal Evaluation of UNEP/GEF project “Enhancing Capacity Knowledge and Technology Support to Build Climate Resilience of Vulnerable Developing Countries”, (GEF 4934), May 2022.
10. Terminal Evaluation of the UNEP/GEF project “Alliance for Zero Extinction (AZE): Conserving Earth’s Most Irreplaceable Sites for Endangered Biodiversity” (GEF 5201), June 2022.
11. Terminal Evaluation of the UNEP/GEF project “LGGE Promoting Energy Efficiency and Renewable Energy in Buildings in Jamaica” (GEF 4167), June 2022.
12. Terminal Evaluation of the UNEP/UNDP/GEF “Building Capacity for LDCs to Participate Effectively in Intergovernmental Climate Change Processes” (GEF 1215), June 2022.
13. Terminal Evaluation of the UNEP Project “GPGC/ENRTP Resource Efficiency Through Application of Lifecycle Thinking” (PIMS 1991), June 2022.
14. Terminal Evaluation of the UNEP Project “EU-UNEP Africa Low Emissions Development Strategies Modelling, Planning and Implementation Project (Africa LEDS Project), June 2022.
15. UNEP/Adaptation Fund project “Enhancing Climate Change Resilience of Rural Communities Living in Protected Areas of Cambodia” (AF Project ID: KHM/MIE/Food/2011/1), July 2022.
16. Terminal Evaluation of the UNEP/GEF project “The GEF Earth Fund: Conservation Agreements Private Partnership Platform” (GEF 4259), September 2022.
17. Terminal Evaluation of the UNEP/GEF project “Energy for Sustainable Development in Caribbean Buildings” (GEF 4171), October 2022.
18. Terminal Evaluation of the UNEP/GEF Project “Applying Landscape and Sustainable Land Management (L-SLM) for Mitigating Land Degradation and Contributing to Poverty Reduction in Rural Areas” (GEF 5825), October 2022.

19. Terminal Evaluation of the UNEP Project “Inquiry into the Design of a Sustainable Finance Systems: Policy Innovations for a Green Economy” (PIMS 01661/623.2) and Terminal Evaluation of the GEF Funded Project “Aligning the Financial System and Infrastructure Investments with Sustainable Development – a Transformational Approach” (GEF 9775), November 2022.
20. Desk Review based on Secondary Data Sources, Conservation and Sustainable Use of Cultivated Tropical Fruit Diversity: Protecting Sustainable Livelihoods, Food Security and Ecosystem Services, (GEF 2430), November 2022.

2023 Evaluation Reports

21. Terminal Evaluation of the UNEP/GEF Project “Increasing Investments in District Energy Systems in Cities – a SE4ALL Energy Efficiency Accelerator” (GEF 9320), December 2022.
22. Terminal Evaluation of the UNEP Project “Haiti Sustainable Energy II (HSE II)” (PIMS 01968), December 2022.
23. Terminal Evaluation of the UNEP/GEF Project “African Rift Geothermal Development Facility (ARGeo), (GEF 2119), January 2023.
24. Terminal Evaluation of the UNEP/GEF Project “The SE4ALL Building Efficiency Accelerator (BEA): Expanding Local Action and Driving National Change” (GEF 9947), January 2023.
25. Terminal Evaluation of the UNEP/GEF Project “Enhanced Cross-Sectoral Land Management Through Land-Use Pressure Reduction and Planning – Serbia” (GEF 5822), by Ms. Justine Braby (Principal Consultant) and Ms. Veronica Muthui (Evaluation Specialist), January 2023.
26. Terminal Evaluation of the UNEP/GEF Project “Promoting Sustainable Land Management (SLM) in Albania through Integrated Restoration of Ecosystems” (GEF 9477), January 2023
27. Terminal Evaluation of the UNEP/GEF Project “Integrating Traditional Crop Genetic Diversity into Technology: Using a Biodiversity Portfolio Approach to Buffer Against Unpredictable Environmental Change in the Nepal Himalayas” (GEF 4464), March 2023.
28. Terminal Evaluation of UNEP/GEF Project “Promotion of Energy Efficiency in Public Lighting in Cote D’Ivoire” (GEF 3876), March 2023
29. Terminal Evaluation of the UNEP/GEF Project “Participatory Sustainable Land Management in the Grassland Plateaus of Western Madagascar (GEF 5354), March 2023.

30. Terminal Evaluation of the UNEP Project “Managing Waste-Water Through Global Partnership” (PIMS 1663), April 2023.
31. Evaluation of the Partnership for Action on Green Economy Interagency Programme ‘Operational Strategy 2016-2020’ (Encompasses UNEP’s PAGE Project, 2018 – 2021) (PIMS 02032), April 2023.
32. Terminal Evaluation of the UNEP/GEF Project “Capacity Building for Information Coordination and Monitoring Systems/SLM in Areas with Water Resource Management Problems of Country Pilot Partnership Program on Sustainable Land Management – Cuba” (GEF 8003), April 2023.
33. Terminal Evaluation of the UNEP/GEF Project “Developing Core Capacity for Decentralized MEA Implementation and Natural Resources Management in Afghanistan” (GEF 5017), May 2023.
34. Terminal Evaluation of the UNEP/GEF Project “Developing Core Capacity for MEA Implementation in Cameroon” (GEF 5060), May 2023.
35. Terminal Evaluation of the UNEP/GEF Project “Capacity Development for the Integration of Global Environmental Commitments into National Policies and Development Decision Making in Bosnia and Herzegovina” (GEF 5302), May 2023.
36. Terminal Evaluation of the UNEP/GEF Project “Increase St. Lucia’s Capacity to Monitor MEA Implementation and Sustainable Development” (GEF 5197), May 2023.
37. Terminal Evaluation of the UNEP/GEF Project “Developing Core Capacity for MEA Implementation in Haiti” (GEF 5557), May 2023.
38. Terminal Evaluation of the UNEP/GEF Project “Scaling Up Sustainable Land Management and Agro-Biodiversity Conservation to Reduce Environmental Degradation in Small-Scale Agriculture in Western Kenya” (GEF 5272), September 2023.
39. Terminal Evaluation of the UNEP/GEF project “Achieving Biodiversity Conservation Through Creation, Effective Management and Spatial Designation of Protected Areas and Capacity Building in Bosnia and Herzegovina” (GEF 6990), October 2023.
40. Desk-Based Study of UNEP Project “Forest Restoration for Enhanced Ecosystems Services Functioning in Afghanistan” (FREESIA) (PIMS 2033), October 2023.
41. Desk-Based Study of the UNEP Project “Opportunities for Mountain Area Integrated Development – Community Conservation and Resilience Initiative” (OMAID) (PIMS 2031), August 2023.

42. Terminal Evaluation of the UNEP/GEF project “Stabilizing GHG Emissions from Road Transport Through Doubling of Global Vehicle Fuel Economy: Regional Implementation of the Global Fuel Economy Initiative (GEF 4909), 2023.
43. Terminal Evaluation of the UNEP/GEF Project “Global Foundations for Reducing Nutrient Enrichment and Oxygen Depletion from Land-based Pollution in Support of Global Nutrient Cycle” (GEF 4212), 2023.
44. Terminal Evaluation of the UNEP Project “Africa’s Coexistence Landscapes: Securing their future for people, elephants and other wildlife” (PIMS 2047), 2023.
45. Terminal Evaluation of the UNEP Project Environment Management Group (2018-2022) (UNEP-PoW 01141), January 2023.
46. Terminal Evaluation of the UNEP/GEF Project “Establishing National Land Use and Land Degradation Profile toward mainstreaming Sustainable Land Management Practices in Sector Policies – SLM Bangladesh ENALULDEP/SLM” (GEF 5823), 2023.

Strategic evaluations / studies

Evaluation of the subprogramme on Environmental Governance, December 2022.

Independent Evaluation of UNEP’s Policy and Strategy for Gender Equality and the Environment (2015), June 2022.

Evaluation of the Partnership for Action on Green Economy Interagency Programme ‘Operational Strategy 2016-2020’ (Encompasses UNEP’s PAGE Project, 2018 – 2021) (PIMS 02032), April 2023

Terminal Evaluation of the UNEP Project “Inquiry into the Design of a Sustainable Finance Systems: Policy Innovations for a Green Economy” (PIMS 01661/623.2) and Terminal Evaluation of the GEF Funded Project “Aligning the Financial System and Infrastructure Investments with Sustainable Development – a Transformational Approach” (GEF 9775), November 2022.

List of projects with validated management-led Terminal Reviews

1. CounterMEASURE II, (PIMS 02049), June 2022.
2. Global Support Programme for Preparation of National Communications and Biennial Update Re-ports of Non-Annex I Parties under the UNFCCC, (GEF 5141), August 2022.

3. Towards coherent implementation of the environmental dimension of the Sustainable Development Goals, (PIMS 02000), October 2022.
4. Equatorial Africa Deposition Network Project, (GEF 3401), July 2023
5. GEMS/Water (Water Quality: Strengthening the normative basis for planning, monitoring, and managing water quality for aquatic ecosystems), (PIMS 02020), August 2022.
6. Strengthening Chile's Nationally Determined Contribution (NDC) Transparency Framework, (GEF 9835), April 2023.
7. Market Policy and legislative development for mainstreaming the sustainable management of marine and coastal ecosystems in Lebanon, (GEF 4020), October 2023.
8. Integrating the Management of Protection and Production Areas for Biodiversity Conservation in the Sierra Tarahumara of Chihuahua, Mexico, (GEF 4883), April 2023.
9. Strengthening national biodiversity and forest carbon stock conservation through landscape-based collaborative management of Cambodia's Protected Areas System as demonstrated in the Eastern Plain Landscape (CAMPAS), (GEF 4905), February 2023.
10. Protecting Biodiversity and Multiple Ecosystem Services in Biological Mountain Corridors in Chile, (GEF 5135), May 2023.
11. Connect: Mainstreaming biodiversity into the heart of government decision making, (GEF 5730), January 2023.
12. Expanding the ongoing support to Least Developed Countries with country-driven processes to advance National Adaptation Plans, (GEF 5868), June 2023.
13. Strengthening Ghana's national capacity for transparency and ambitious climate reporting, (GEF 9820), September 2023.
14. EUROCLIMA+ Accelerating Climate Action in Latin America, (PIMS 02006), June 2023.
15. Improving Ecosystem Management for Sustainable Livelihoods within the Framework of Lancang Mekong Cooperation, (PIMS 2101), October 2023.

Annex 2

List of evaluations assessed for responsiveness to human rights and gender equality, 2022-2023

Projects that scored a Highly Satisfactory rating

1. Conservation and Sustainable Use of Cultivated Tropical Fruit Diversity: Protecting Sustainable Livelihoods, Food Security and Ecosystem Services, (GEF 2430) November 2022
2. Opportunities for Mountain Area Integrated Development – Community Conservation and Resilience Initiative (OMAID) – PIMS 2031

Projects that scored a Satisfactory rating

1. Conservation Agreement Private Partnership Platform GEF 4259
2. Alliance for Zero Extinction (AZE): Conserving Earth’s Most Irreplaceable Sites for Endangered Biodiversity” GEF ID 5201 Alliance for Zero Extinction (AZE): Conserving Earth’s Most Irreplaceable Sites for Endangered Biodiversity GEF ID 5201

3. Enhancing Climate Change Resilience of Rural Communities Living in Protected Areas of Cambodia
4. Support to Eligible Parties to Produce the Sixth National Report to the CBD – (Global: Africa-3, Maldives, Nicaragua, Pakistan and Solomon Islands) GEF ID: 9832
5. Support to Eligible Parties to Produce the Sixth National Report to the CBD (Africa-1) GEF ID: 9817
6. Support to Eligible Parties to Produce the Sixth National Report to the CBD (Africa-2) GEF ID: 9824
7. Support to Eligible Parties to Produce the Sixth National Report to the CBD (Europe, CIS and Mongolia) GEF ID: 9822
8. Support to Eligible Parties to Produce the Sixth National Report to the CBD (Pacific) GEF ID: 9823
9. African Rift Geothermal Development Facility (ARGeo), GEF 2119
10. The SE4ALL Building Efficiency Accelerator (BEA): Expanding Local Action and Driving National Change” (GEF ID 9947)
11. Enhanced Cross-Sectoral Land Management Through Land-Use Pressure Reduction and Planning – Serbia GEF 5822
12. Promoting Sustainable Land Management (SLM) in Albania through Integrated Restoration of Ecosystems (GEF 9477)
13. Participatory Sustainable Land Management in the Grassland Plateaus of Western Madagascar (GEF 5354)
14. Managing Waste-Water Through Global Partnership – PIMS 1663
15. Capacity Building for Information Coordination and Monitoring Systems/ SLM in Areas with Water Resource Management Problems of Country Pilot Partnership Program on Sustainable Land Management - Cuba (GEF 8003)
16. Developing Core Capacity for Decentralized MEA Implementation and Natural Resources Management in Afghanistan – GEF 5017,
17. Forest Restoration for Enhanced Ecosystems Services Functioning in Afghanistan” (FREESIA) – PIMS 2033
18. Environment Management Group (2018-2022) (UNEP-PoW 01141)
19. Establishing National Land Use and Land Degradation Profile toward mainstreaming Sustainable Land Management Practices in Sector Policies – SLM Bangladesh ENALULDEP/SLM (GEF 5823)

Projects that scored a ‘Moderately Satisfactory’ rating

1. UNEP Inquiry into the Design of a Sustainable Financial System (PIMS 01661)
2. Building Capacity for LDCs to participate effectively in intergovernmental climate change processes.” GEF ID 1215
3. Resource Efficiency through Application of Lifecycle thinking” (REAL) - PIMS 1991
4. Enhancing Capacity, Knowledge and Technology Support to Build Climate Resilience of Vulnerable Developing Countries / Ecosystem-based Adaptation through South-South Cooperation (EbA South) GEF 4934
5. Enabling South Africa to Prepare its Third National Communication (TNC) and Biennial Update Report (BUR-2) to the UNFCCC Enabling South Africa to Prepare its Third National Communication (TNC) and Biennial Update Report (BUR-2) to the UNFCCC GEF 5237
6. Applying Landscape and Sustainable Land Management for mitigating land degradation and contributing to poverty reduction in rural areas GEF 5825
7. Increasing Investments in District Energy Systems in Cities – a SE4ALL Energy Efficiency Accelerator” (GEF ID 9320)
8. Integrating Traditional Crop Genetic Diversity into Technology: Using a Biodiversity Portfolio Approach to Buffer Against Unpredictable Environmental Change in the Nepal Himalayas GEF 4464
9. Promotion of Energy Efficiency in Public Lighting in Cote D’Ivoire GEF 3876
10. Developing Core Capacity for MEA Implementation in Cameroon – GEF 5060
11. Capacity Development for the Integration of Global Environmental Commitments into National Policies and Development Decision Making in Bosnia and Herzegovina – GEF 5302
12. Increase St. Lucia’s Capacity to Monitor MEA Implementation and Sustainable Development – GEF 5197
13. Developing Core Capacity for MEA Implementation in Haiti – GEF 5557
14. Achieving Biodiversity Conservation Through Creation, Effective Management and Spatial Designation of Protected Areas and Capacity Building in Bosnia and Herzegovina (GEF 6990)
15. Stabilizing GHG Emissions from Road Transport Through Doubling of Global Vehicle Fuel Economy: Regional Implementation of the Global Fuel Economy Initiative (GFEI)- GEF 4909

Projects that scored ratings in the unsatisfactory range

16. EU-UNEP Africa Low Emissions Development Strategies Modelling, Planning and Implementation Project (Africa LEDS project)
17. Energy for Sustainable Development in Caribbean Buildings GEF ID 4171
18. Haiti Sustainable Energy II (HSE II) - PIMS 01968
19. Scaling Up Sustainable Land Management and Agro-Biodiversity Conservation to Reduce Environmental Degradation in Small-Scale Agriculture in Western Kenya (GEF 5272)
20. Africa's Coexistence Landscapes: Securing their future for people, elephants and other wildlife (PIMS ID 2047)
21. Mainstreaming Biodiversity Conservation and Sustainable Use for Improved Human Nutrition and Well-being (BFN) GEF ID: 3808
22. Mainstreaming Agrobiodiversity Conservation and Use in Sri Lankan Agro-Ecosystems for Livelihoods and Adaptation to Climate Change" - GEF ID Number 4150
23. LGGE Promoting Energy Efficiency and Renewable Energy in Buildings in Jamaica, GEF ID: 4167
24. Partnership for Action on Green Economy Interagency Programme 'Operational Strategy 2016-2020' PIMS 02032

Annex 3

UNEP's results definitions

The UNEP Evaluation Office and the Project Review and Oversight Unit of the Policy and Programme Division (PPD) have, over several years worked together to develop and periodically refine a glossary of UNEP results definitions³³. The application of these definitions informs project design implementation and evaluation. A summary of those terms most closely associated with evaluation processes is shown below in **Table 7**. Several important distinctions continue to need emphasis in the application of Results-based Management principles:

Activity - Deliverable - Output

The work on results' definitions provided an opportunity to clarify that **activities** (work performed) and **deliverables** (products or services derived from a completed activity) differ from **Outputs** (immediate gains experienced from the work performed and products or services delivered) by the inclusion of the intended beneficiary or user. From a results perspective, the attention is placed on those

³³ The UNEP results definitions build on those of OECD-DAC, UNDA, UNDP, UNEG, and the World Bank.

who are affected by projects rather than those who are responsible for delivering them. Importantly, this refinement in the definition of an output shifts the perspective from that of the supplier (what has been delivered) to that of the beneficiary (what has been gained); it does not change the ambition underpinning the output, nor does it change the timeframe within which Outputs are made available. For example, in this way the ‘provision of technical assistance’ (activity) builds the knowledge and understanding of a target audience (output), which may lead to a change of behaviour, based on that new knowledge and understanding, by the target audience (outcome).

Project Direct Outcome - Project Outcome

Evaluations aim to assess what has been achieved against what was planned (and committed to). To clearly represent the ambition of, and commitment made by, a project the term **Project Outcome** is used by UNEP to refer to those Outcomes a project is expected to achieve by the time it is operationally complete. However, it is recognised that some Outcomes (i.e., the use or uptake of Outputs) occur before the end of the project and these are referred to, in all parts of the house, as **Project Direct Outcomes**. Project Outcomes are the highest level of result to which UNEP evaluations hold projects to account.

Attribution - Contribution - Credible Association

The gold standard of proving that a project has achieved its results is demonstrating attribution between the observed effects and the efforts of the project. However, this is not always possible, and it is helpful to consider a wider range of ways of establishing claims between project efforts and observed results.

In order to **attribute** any Outcomes and impacts to a project intervention, one needs to consider the difference between what has happened *with*, and what would have happened *without*, the project (i.e., take account of changes over time and between contexts in order to isolate the effects of a project). This requires appropriate baseline data and the identification of a relevant counterfactual, both of which are frequently not available for evaluations.

Establishing the **contribution** made by a project in a complex change process relies heavily on prior *intentionality* (e.g., approved project design documentation, logical framework) and the articulation of *causality* (e.g., narrative and/or illustration of the Theory of Change). Robust evidence that a project was delivered as intended and that the expected causal pathways developed as designed, can be used to support claims of contribution. Such claims can be strengthened where an alternative theory of change can be excluded (e.g., that the ‘business as usual’ behaviour pattern is no longer evident or prevalent).

Finally, a **credible association** between the implementation of a project and observed positive effects can be made where a strong causal narrative, although not explicitly articulated, can be inferred by the chronological sequence of events, active involvement of key actors and engagement in critical processes.

Programme - Portfolio

There has been a tendency, both in UNEP and elsewhere, for these two terms to be used either interchangeably or at least, loosely. There is, however, an important distinction between the two. Whereas the projects within both a portfolio and a programme will share certain common characteristics with each other, only those projects in a programme will be working towards common Outcomes and be managed in a coordinated way that generates additional benefits not available when implemented separately.

Evaluations of portfolios can provide insights on key issues based on multiple experiences of working on a common theme or in a common geographic area etc., but evaluations of programmes can provide evidence of aggregated and synergistic effects across a group of projects.

Table 7. UNEP results definitions relevant to evaluations

TERM	UNEP results definitions - relevant to evaluations
Activity	An action taken, or work performed, through which inputs are utilized to realize specific results
Assumption	An assumption is a significant external factor or condition that needs to be present for the realization of the intended results but is beyond the influence of the project and its partners. Assumptions are often positively formulated risks. (See also Driver)
Attribution	Attribution can be claimed when comprehensive evidence proving the cause-and-effect relationship between the project and the observed results is presented. To make a strong claim of attribution one needs to be able to isolate the effects of an intervention from changes over time and differences in contexts (i.e. have an M&E design that provides data on indicators before and after project implementation and a robust counterfactual).
Contribution	Contribution can be claimed when compelling evidence is presented that supports a cause-and-effect relationship through which intended collective results are achieved by the combined efforts of more than one project.

TERM	UNEP results definitions - relevant to evaluations
Credible association	Credible association can be claimed when compelling evidence supporting a cause-and-effect relationship between the project and the observed results is present. A claim to a credible association between project and results can be made based on 'intentionality' (normally reflected in a project design document), the articulation of causality (ideally through a detailed/nuanced TOC) and evidence (through the chronology of events, action of identified key stakeholders etc) that the intention was followed, and the expected causality pathways emerged.
Driver	A driver is a significant external factor that, if present, is expected to contribute to the realization of the intended results of a project. Drivers can be influenced by the project and its partners. (See also Assumption).
Direct outcome(s)	A direct outcome is an outcome that is intended to be achieved from the uptake of Outputs and occurring prior to the achievement of Project Outcome(s).
Deliverable	A deliverable is a specified product or service derived from completed activity(ies). For example, a draft report, a revised manual, X number of workshop facilitation days are all frequent deliverables in a UNEP project context. (Deliverables are viewed from the perspective of the supplier/provider of the product or service).
Effect	An effect is a change which is a consequence of an action or other social, economic, political or environmental cause. These changes can be intended, unintended, positive or negative.
Evaluand	The evaluand is the entity being evaluated.
Evaluand parameters	An evaluand's parameters are the features of an evaluand that define its scope (e.g. timeframe, funding envelope, results framework, geographic dimensions).
Evaluation	Evaluation is the systematic and objective assessment of an on-going or completed project, programme, strategy or policy (i.e. 'entities'), its design, implementation, results and likelihood of impact
Expected accomplishments	Expected Accomplishments are the Outcomes of a UNEP Programme of Work (i.e. the Outcomes specified in a Sub-Programme).
Goals	Goals are the higher-order result(s) to which a project is intended to contribute.

TERM	UNEP results definitions - relevant to evaluations
Impact	Impacts are long-lasting results arising, directly or indirectly from a project. Impacts are intended and positive changes and must relate to UNEP's mandate.
Impact pathway / Causal pathway / Results chain	Impact or Causal Pathways and Results Chains all describe cause and effect relationships between Outputs, Outcomes and impacts and are the basis of a programme/project's "Theory of Change".
Intermediate states	Intermediate states are changes (i.e. changes at the outcome level) beyond the Project Outcome(s) that are required to contribute towards the achievement of the intended impact of a project.
Lessons learned	Lessons learned reflect the new knowledge or understanding gained by the experience of implementing a project that is applicable to, and useful in, other similar contexts
Logical framework	A Logical Framework (or logframe) is a tool for summarizing a project or programme's intended results. It specifies project results, indicators and their baseline and target values. It also includes a milestone schedule to deliver the expected output(s) and/or achieve intended result(s). The format or content of a logframe may vary from one organization to the other.
Mid-term Evaluation	A Mid-Term Evaluation is a formative evaluation performed towards the middle of the period of implementation of a project.
Monitoring	Monitoring is a continuing function that uses the systematic collection of data on project / programme implementation (e.g. completion of activities, rate of expenditure, emergence of risks, milestone delivery, inclusive participation of intended stakeholders, etc.) to provide management with indications of the extent of progress against plans and targets.
Objectives	An objective describes the overall intention, which should be achieved at the end of a defined period. (In UNEP, the Project Objective is not considered a higher-order result and does not appear in the Logical Framework or Theory of Change).
Outcome(s)	An outcome is the use (i.e., uptake, adoption, application) of an output by intended beneficiaries, observed as a change in institutions or behaviours, attitudes or conditions

TERM	UNEP results definitions - relevant to evaluations
Outputs	An output is the availability (for intended beneficiaries/users) of new products and services and/or gains in knowledge, abilities and awareness of individuals or within institutions. For example, access by the intended user to a report; new knowledge held by a workshop participant at the end of a training event; heightened awareness of a serious risk among targeted decision-makers. (Outputs are viewed from the perspective of the intended beneficiary or user of the output rather than the provider).
Performance	Performance is the degree to which a project is implemented, or a partner operates, according to specific criteria / standards/ guidelines or achieves intended results.
Portfolio	A group of projects and/or programmes that share a common characteristic relevant to the organization's strategic objectives (e.g. funded by the same donor, operating in the same country, thematic area etc).
Programme	A programme is a group of synergistic projects contributing to a common outcome(s) and managed in a coordinated way to obtain benefits not available from managing the projects individually.
Project	A project is a time-bound intervention with a specific funding envelope that addresses a defined set of results within an identified implementation context or geographic area ³ . The main components of the project must be interlinked/ interdependent to achieve the project outcome(s).
Project components	A descriptive label related to the way in which work is planned and managed. A Project Component can define a stream of work
Project outcome(s)	Project Outcome(s) are those Outcomes that are intended to be achieved by the end of project timeframe/funding envelope.
Results	Results are intended changes in a state or condition that derive from a cause-and-effect relationship. Such changes must be describable and measurable/discernible. A results statement, the associated targets and their indicators should conform collectively to the SMART ⁴ and/or CREAM ⁵ principles. Outputs, Outcomes and impact are considered 'results' (as opposed to inputs and activities).
Safeguard	Safeguards are a management approaches taken to avoid negative unintended effects of a project

TERM	UNEP results definitions - relevant to evaluations
Sustainability	Sustainability is the continuation of project Outcomes and their associated benefits after the completion of a project
Terminal Evaluation	A Terminal Evaluation is a summative evaluation performed at the operational completion of a project or programme
Theory of Change	A Theory of Change is a method used for planning a project, describing the participation that will be needed by different actors and for evaluating the project's performance. It articulates long lasting intended impact and then maps backward to identify the preconditions necessary to achieve this impact(s). It is a comprehensive description and illustration of how and why a desired change is expected to happen in a context. A Theory of Change also allows for unintended positive and/or negative effects to be depicted

Annex 4

Weightings for evaluation criteria

Table 8. Weightings applied across evaluation criteria for overall performance ratings

Criterion	Sub-criteria/ factor weight	Criterion Weightings
A. Strategic Relevance		6
1. Alignment to UNEP's MTS, POW and strategic priorities	0.5	
2. Alignment to Donor/Partner strategic priorities	0.5	
3. Relevance to regional, sub-regional & national environmental priorities	2.5	
4. Complementarity with relevant existing interventions	2.5	
B. Quality of Project Design		4
C. Nature of External Context *		
D. Effectiveness		45
1. Availability of Outputs	5	
2. Achievement of project Outcomes	30	
3. Likelihood of impact	10	

Criterion	Sub-criteria/ factor weight	Criterion Weightings
E. Financial Management		5
1. Adherence to UNEP's policies and procedures	1.66	
2. Completeness of project financial information	1.66	
3. Communication between finance and project management staff	1.66	
F. Efficiency		10
G. Monitoring and Reporting		5
1. Monitoring design and budgeting	1.66	
2. Monitoring of project implementation	1.66	
3. Project reporting	1.66	
H. Sustainability		20
1. Socio-political sustainability	6.66	
2. Financial sustainability	6.66	
3. Institutional sustainability	6.66	
I. Factors Affecting Performance / Cross-Cutting Issues		5
1. Preparation and readiness	0.71	
2. Quality of project management and supervision	0.71	
3. Stakeholders participation and cooperation	0.71	
4. Responsiveness to human rights and gender equality	0.71	
5. Environmental and social safeguards	0.71	
6. Country ownership and driven-ness	0.71	
7. Communication and public awareness	0.71	
Overall Project Rating		100

*The criterion 'Nature of the external context' is not used directly for assessing overall project performance. It is used indirectly in situations where operational conditions are especially challenging.

