United Nations E/ICEF/2026/17

Economic and Social Council

Distr.: General
15 May 2026

Original: English

For decision

United Nations Children's Fund
Executive Board
Annual session 2026
16–19 June 2026
Item 11 of the provisional agenda*

Annual report for 2025 on the evaluation function in UNICEF

Summary

This report provides an overview of the UNICEF evaluation function in 2025. It provides a contextual overview, highlights key achievements and challenges, and concludes with reflections on priorities for further strengthening the function and its enabling environment in the years ahead.

Elements of a draft decision for consideration by the Executive Board are provided in section V.

* E/ICEF/2026/9.
Note: The present document was processed in its entirety by UNICEF.

26-06992 (E) 180526

I. Overview

1. The year 2025 represented a crossroads: the close of the 2022–2025 quadrennium, the second full year of implementation of the Revised evaluation policy of UNICEF (E/ICEF/2023/27 and E/ICEF/2023/27/Corr.1), and an unprecedented period of organizational strain. The Future Focus Initiative (FFI) and its associated Programme Budget Review introduced significant restructuring and fiscal tightening, placing the evaluation function under pressure to adapt while continuing to deliver – and while safeguarding the independence that underpins its credibility.

2. Against this backdrop, performance on the function's key indicators remained broadly stable – a result that was neither inevitable nor easily won. Evaluation submissions increased compared with 2024, reflecting deliberate efforts to strengthen coverage alongside programme cycle requirements, while remaining below earlier quadrennium peaks, in keeping with the function's sustained emphasis on quality and strategic relevance over volume. Evaluation quality held steady, with a slight improvement in mean scores in 2025. Timeliness of management responses showed gradual improvement, though a significant share continued to exceed the 60-day threshold – a persistent challenge that reflects the decentralized nature of the function and the limits of what evaluation alone can drive. Evaluation expenditure remained stable in absolute terms, but consistently below the 1 per cent benchmark, pointing to a constrained and unevenly financed enabling environment.

3. While quantitative measurement of evaluation's collective contribution to organizational decision-making remains an ongoing challenge, the qualitative evidence of impact continued to grow. The 2025 Study of Influential Evaluations provided systematic analysis of how and why evaluations make a difference – identifying recurring patterns of influence across institutional reform, programme design, national policy and humanitarian response. Evaluations, syntheses and advisory notes informed key corporate decisions, including preparation of the UNICEF Strategic Plan, 2026–2029 and the organization's navigation of the FFI process itself. These examples illustrate a function whose influence extends well beyond what performance indicators alone can capture.

4. Three cross-cutting strategic initiatives – impact evaluation, national evaluation capacity development and global evaluation synthesis – collectively advanced the function's most ambitious objective: helping UNICEF and its partners focus on what works and achieve impact-level results for children at scale. By generating rigorous evidence on effective interventions, building the national systems through which Governments hold themselves accountable for results, and translating accumulated knowledge into decision-ready insights for high-level policy forums, these initiatives extended the reach and influence of evaluation well beyond what UNICEF alone can achieve.

5. Internally, the function made significant strides in both the hard-wiring and soft-wiring of a maturing evaluation system. The first-ever *UNICEF Evaluation Handbook* was completed, risk-based planning tools were institutionalized, and a suite of methodological guidance was developed. The Global Evaluation Leadership Team (GELT) crystallized as a vehicle for coherence and collective stewardship across the function's three levels. And the Plan for global evaluations, 2026–2029 – a portfolio of 23 evaluations, compared with 44 in the previous quadrennium, fully costed and accompanied by the first-ever costed annual workplan for the Evaluation Office – was endorsed by the Executive Board in February 2026, reflecting deliberate, risk-based prioritization, and significantly strengthening transparency in evaluation planning and resourcing.

6. The evaluation function enters the new quadrennium with reduced resources and a changed institutional landscape, but also with stronger foundations than at any previous point in its history. Closing the estimated $7.9 million funding gap in the Plan for global evaluations, 2026–2029, maintaining the institutional safeguards for independence, and sustaining the momentum built in the areas of quality, utilization and strategic leverage will be the central challenges ahead. The forthcoming independent external assessment of evaluation in UNICEF, to be conducted by the Organisation for Economic Co-operation and Development-Development Assistance Committee (OECD-DAC) and presented to the Executive Board in 2027, will provide a timely and fully independent perspective on the health and maturity of the function – and an important opportunity to chart the course for the quadrennium ahead.

II. Key developments in the evaluation function

A. Background and priorities for 2025

Uncertainty and change in the final year of the 2022–2025 quadrennium

7. As the final year of the 2022–2025 quadrennium, 2025 represented an important juncture: the end stage of the implementation of the Plan for global evaluations, 2022–2025 and the end of first two years of the implementation of the Revised evaluation policy of UNICEF (2023). The quadrennium had witnessed significant challenges, described in annual reports of the evaluation function presented to the Executive Board during this period, and the Revised evaluation policy had signalled important strategic shifts to reposition the evaluation function to address these challenges and optimize the function's value in supporting organizational learning and accountability for the years to come. The priorities for the year reflected the ongoing effort to implement the policy and monitor the key performance indicators associated with it, while also contributing to the development of the UNICEF Strategic Plan, 2026–2029.

8. This approach of sustaining the momentum built over the quadrennium was soon overtaken by an increasingly challenging operating environment for UNICEF. The FFI that ensued introduced significant organizational reform and financial uncertainty. For the evaluation function, FFI led to a sharp reduction in the Evaluation Office resource envelope for 2026–2029, exceeding the 25 per cent organization-wide benchmark, and a relocation of the office to Rome. Resources available through the Evaluation Pooled Fund were also reduced, and governance arrangements were modified, including the role of evaluation leadership in assessing funding needs. Implications on staffing were also significant. Evaluation posts were abolished, reassigned, or shifted from institutional resources to other resources funding, reducing capacity below minimum policy-mandated staffing levels. These developments placed pressure on coverage, capacity and resourcing across the function, with reduced resource envelopes meaning that some evaluations planned for 2025 and the next quadrennium could not be fully funded or required delayed implementation.

9. FFI thus defined the operating context for 2025 in ways that were both constraining and clarifying. It required the evaluation function to absorb significant cuts while simultaneously maintaining delivery, protecting independence and supporting an organization in the midst of unprecedented change – a combination of demands that made prioritization and adaptability the defining imperatives of the year.

Recalibrating priorities in an evolving operating environment

10. In response to the wider organizational context of uncertainty and change, the evaluation function pivoted towards four parallel priorities in 2025: (1) adapting the function's own structure and operations to the emerging realities while simultaneously

accompanying the organization through an unprecedented period of institutional strain, and while safeguarding its independence; (2) continuing to deliver on its core commitment to produce timely, relevant, credible and useful evaluations that add value to UNICEF and its partners; (3) reinforcing the momentum it had built to "move the needle" by strengthening evaluation practice internally and shaping global evidence systems for children externally; and (4) preventing backsliding on the key performance indicators established for the function. Navigating these priorities simultaneously, under significant resource pressure and organizational uncertainty, defined the work of the function throughout the year.

11. With respect to the first priority, the function was not only downsized; rather, it also sought to adapt to increasing financial constraint by accelerating prioritization- and efficiency-enhancing measures, some of which had already been planned or were under way prior to FFI. These included the selective introduction of a hybrid delivery model, whereby in-house evaluation expertise is blended with individual consultancies as an alternative to often-costlier institutional consultancies. They also included changes to business processes that improved timeliness and sharpened focus on substance and methodology while also lowering cost, as well as a restructuring of a reduced Evaluation Office staff complement to maximize coherence and cohesion within the team. Creative and innovative methodological and technological solutions, described in greater detail in section IIC below, further demonstrated that cost-efficiency can in some cases mutually reinforce evaluation timeliness, relevance quality and credibility.

12. Prioritization of evaluations became even more essential as a means of ensuring that increasingly finite resources were directed towards areas of highest strategic importance – and, accordingly, highest potential value in supporting learning and accountability. This shift was reflected in the Plan for global evaluations, 2026–2029, developed for presentation in draft form to the Executive Board at its second regular session of 2025 and endorsed at the first regular session of 2026. The Plan introduced a risk- and opportunity-based approach to selecting priorities and a structured tiering of evaluations according to their criticality level. The resulting portfolio of 23 evaluations – a significant reduction from the 44 evaluations planned in the previous quadrennium – also demonstrated heightened sensitivity to corporate absorptive capacity constraints. The first-ever presentation of a costed annual workplan for 2026, which had been requested by the Executive Board, significantly strengthened transparency in relation to evaluation resourcing – and to the gap between resources and critical evaluation priorities.

13. While adapting itself to a changing operating context and declining resources, the evaluation function actively sought to support the organization in navigating FFI with evidence-based analysis. From the outset of the reform process, the Evaluation Office prepared two advisory notes, one of these together with the Office of Internal Audit and Investigations, that drew on evaluations, audits and strategic reviews since 2017 to highlight recurring challenges and lessons relevant for decision-making. Members of the evaluation function also participated in discussions with regional management teams with the aim of supporting adherence to policy commitments, safeguard regional capacity, and continuing to cultivate the strong relationships that underpin evaluation success.

14. The independence of an evaluation function hinges on adequate and predictable resources, freedom to pursue its work without reprisal, and inclusion in decisions and actions that directly affect it. Attending to these hallmarks of independence in the face of unprecedented organizational change and financial constraint therefore constituted another key priority in 2025. Executive Board decision 2025/14 reaffirmed the independence of the evaluation function and requested preparation of costed annual

evaluation workplans beginning in 2026. In late 2025, moreover, preliminary measures were taken to prepare for an independent external assessment of evaluation in UNICEF. This assessment, which was requested by the Executive Board in its decisions 2023/12 and 2023/20, will provide UNICEF management and the Executive Board with an independent, impartial perspective on the health and maturity of the evaluation function and its enabling environment, including independence-related measures.

15. Taken together, these efforts associated with the first strategic priority aimed to reinforce the institutional foundations of the evaluation function and position it to operate as effectively and efficiently as possible in a future UNICEF. The sections that follow document progress on the other three priorities: the evaluations and evidence that continued to deliver value for children and the organization (section IIB); the innovations and methodological advances that extended the function's reach under constrained conditions (section IIC); and the strategic initiatives through which evaluation is helping to build a broader movement of evidence-based accountability for children (section IID).

B. Noteworthy evaluations contributing to organizational results

16. A commitment to understanding and demonstrating how evaluation makes a difference has been a hallmark of the function throughout the quadrennium. In 2025, this commitment produced a major milestone, the *Study of Influential Evaluations*,[1] which examined how evaluation findings translate into concrete decisions within UNICEF and external policy environments. The study identified four recurring pathways through which evaluation contributes to better results for children: informing institutional reform and strategic positioning; guiding programme redesign and scale decisions; shaping national policy and systems development; and strengthening accountability and operational learning in fragile and high-risk contexts. The examples below illustrate each of these pathways.

Informing institutional reform and strategic positioning

17. An *Evaluation of Human Resources Management in UNICEF* found progress in recruitment, gender parity and workforce diversity, supported by digital human resources systems and staff well-being initiatives while identifying gaps in strategic workforce planning, career development and organizational culture. A cultural "thumbprint" analysis showed alignment with UNICEF core values, but highlighted challenges related to risk aversion and hierarchical decision-making, informing reforms to workforce planning, recruitment and performance management. A *Baseline Assessment of UNICEF Human Resources Safeguarding Practices* identified gaps in safeguarding integration across the human resources lifecycle and established a road map for embedding safeguarding in recruitment, training and performance systems. An *Evaluation of UNICEF Innovation*, for its part, found progress in governance, portfolio management and partnerships, while highlighting fragmented oversight and limited evidence on the results of scaled innovations, informing improvements in innovation governance and financial tracking. Strong engagement between the Evaluation Office and the Division of People and Culture and the Office of Innovation has resulted in strong uptake of all three evaluations.

18. Sector-focused evaluations also contributed to institutional positioning. The *Evaluation of UNICEF's Work in Primary Healthcare* found that more than 80 per

[1] UNICEF, The 2025 Study of Influential Evaluations*: Final report*, UNICEF, New York, 2025, available at https://www.unicef.org/evaluation/media/7386/file/The%202025%20Study%20of%20Influential%20Evaluations.pdf.pdf.

cent of country offices contribute to primary healthcare-related interventions, with strongest results where service delivery was combined with upstream engagement on policy, governance and financing. These findings are informing efforts to strengthen cross-sector integration and clarify the UNICEF corporate approach to primary healthcare. At regional level, the *Multi-Country Evaluation of Immunization Programming in the Europe and Central Asia Region* identified opportunities to strengthen digitalization, outreach and regional collaboration for sustained immunization coverage.

Guiding programme design and scale decisions

19. The *Evaluation of the Accelerating Sanitation and Water for All Programme* found that it substantially expanded access to improved water, sanitation and hygiene (WASH) services across eight countries, including in schools, health facilities and communities, while identifying sustainability challenges. The management response strengthened monitoring frameworks, cross-sector collaboration and sustainability checks, informing the design and financing of the programme's next phase.

20. Similarly, the *Evaluation of UNICEF work in early childhood development and early childhood education, 2018–2023* (E/ICEF/2026/3) examined the contribution of UNICEF to strengthening early childhood systems, confirming its strong role in supporting Governments to develop multisectoral policies and plans. The evaluation also highlighted the contribution of UNICEF to strengthening government capacity to mobilize financing for early childhood programmes, while noting that public financing strategies have not yet translated into large-scale investments in early childhood services. This evaluation's recommendations are informing updates to guidance, indicators and results chains under the Strategic Plan cycle, as indicated at the first regular session of 2026.

21. Evidence generated at country level also informed programme adjustments. The *Evaluation of the Sustainable WASH for Resilience Programme in South Sudan* identified sustainability gaps while confirming progress in expanding safe water and sanitation services in fragile contexts, informing adjustments to programme design and strengthening collaboration with national and community actors responsible for maintaining water systems. The *Evaluation of the Nigeria-UNICEF WASH Country Programme Cooperation (2018–2022)* documented progress in expanding access to WASH services while identifying opportunities to strengthen coordination with government systems and explore innovative financing approaches, informing planning for the next programme phase.

22. The *Evaluation of Resilience Integrated Education Programming for Children in North-East Nigeria* and the *Evaluation of the Education Cannot Wait Multi-Year Resilience Programme II*, meanwhile, generated evidence on resilience-oriented education programming in conflict-affected contexts, informing improvements in teacher training, community engagement and integration of education interventions with broader resilience approaches.

Shaping national policy and systems development

23. The *Evaluation of the Integrated Strategy on Chronic Malnutrition in Guatemala* contributed evidence to national policy dialogue and strengthened coordination mechanisms on chronic child malnutrition. The *Evaluation of the National Plan of Action for Nutrition in Malaysia* informed alignment of sector strategies and financing arrangements. The *Evaluation of the Territorial Strategy for the Prevention and Reduction of Chronic Child Undernutrition in Ecuador* generated evidence to strengthen implementation of national nutrition policies at subnational level. At the municipal level, the *Evaluation of the Municipality United for Children*

and Adolescents Programme in Argentina showed how UNICEF supports strengthened local planning and coordination for child-focused policies and services, informing the redesign of the subsequent country programme.

Strengthening accountability and operational learning in fragile contexts

24. The *Evaluation of UNICEF's L3 Response in Sudan* highlighted the role of UNICEF in sustaining humanitarian operations, while identifying limitations related to preparedness, early warning and staffing gaps. The *Evaluation of UNICEF's L3 Response in Haiti* emphasized the need for stronger preparedness and more integrated humanitarian strategies supporting essential services for children. The *Global evaluation of UNICEF emergency preparedness in humanitarian action* also identified progress in strengthening preparedness systems while highlighting persistent gaps such as underfunded preparedness functions and inconsistent risk analysis.

25. At the decentralized level, the *Evaluation of the UNICEF Response to the Refugee Crisis in Eastern Chad* informed adjustments to humanitarian coordination supporting refugee children and host communities, while the *Evaluation of the Joint Resilience Programme in South-Central Somalia* demonstrated how evidence on resilience programming can strengthen planning of protection and social protection interventions supporting food security, education and protection outcomes for children in fragile contexts.

26. The *Study of Influential Evaluations* identified enabling conditions that strengthen evaluation influence. In a context of reform and fiscal tightening, these conditions enabled evaluation evidence to guide prioritization and decision-making. The examples documented in this section reflect precisely that: timely, decision-oriented evaluations that strengthened programmes, systems and policies for children across a wide range of contexts.

C. Moving the needle internally: Strengthening the "hard-wiring" and "soft-wiring" of the function

27. As declining resources coincided with evidence pointing to areas in need of strengthening, initiatives launched earlier in the quadrennium to enhance the relevance, credibility, utility and independence of evaluations took on even greater urgency with the onset of the FFI. The Revised evaluation policy (2023) had categorized these measures as the "hard-wiring" and "soft-wiring" of a robust evaluation function. Progress was significant in both areas in 2025.

28. "Hard-wiring" refers to the concrete outputs that operationalize the policy in more detailed guidance and procedures. A landmark achievement was the drafting of the first-ever *UNICEF Evaluation Handbook*, which consolidates methodological standards and operational guidance for the function as a whole. Risk-based planning approaches were institutionalized, culminating in a far more data-driven, evidence-based Plan for global evaluations, 2026–2029 than ever before – one developed through a bespoke platform for identifying high-criticality evaluation themes. Guidance for costed evaluation plans was refined; guidance notes on gender equality integration and youth engagement were developed; a suite of advanced methodological guidance notes was produced; quality assurance tools were updated; a template for assessing impact evaluations was introduced; and the evaluation policy implementation procedure progressed towards institutional clearance. Preparatory work also began for the independent external assessment of evaluation in UNICEF. Progress on several of these touchstone resources will continue in 2026.

29. "Soft-wiring" refers to the relationships – within the function itself, and between the evaluation function and its partners elsewhere in UNICEF – that are needed to internalize the principles articulated in the evaluation policy. Within the function itself, this element was particularly important for a function that had only been constituted as a global (i.e. centralized and decentralized) function in 2019, had experienced a leadership gap in 2020–2021, and had emerged from remote operations during the COVID-19 pandemic in 2022 – and one that operates in a matrix management arrangement at the decentralized level. Soft-wiring within the function was strengthened through the formalization of GELT, consisting of Evaluation Office management and Regional Evaluation Advisers, to reinforce coherence in planning, policy implementation and strategic positioning, including development of the Plan for global evaluations. Direct and ad hoc engagement with Regional Evaluation Advisers also deepened, strengthening alignment, accountability and collective stewardship of the function, and standards for independence at the decentralized level.

30. Beyond its progress in advancing these policy commitments, the function put in place numerous other measures to strengthen how the function plans, conducts and communicates about evaluations. The risk-based approach to producing a highly prioritized, strategically aligned Plan for global evaluations, 2026–2029 was one such advance, entailing the development of a bespoke data-driven platform for identifying high-criticality evaluation themes.

31. In the conduct of its evaluations, the function sought creative and innovative solutions to complex evaluation challenges. The *Evaluation of Human Resources Management in UNICEF* applied new ways of harnessing case studies and data visualization tools to analyse workforce trends and organizational culture. Similarly, the *Evaluation of Innovation* examined how changes across governance arrangements, organizational systems and partnerships influence the scaling of innovative solutions, drawing on case studies from multiple regions. These analyses helped identify institutional barriers and enabling factors affecting performance.

32. Adaptive methodological approaches proved particularly useful in contexts where access and security constraints limited traditional fieldwork. During the evaluation of the *Level 3 response in the State of Palestine*, for example, an early inception mission in Gaza enabled rapid feedback supporting operational adjustments. In other fragile settings, such as Fiji, Lebanon, Liberia and Türkiye, phased and adaptive approaches were applied to maintain analytical coherence despite operational disruptions. In several regions, evaluations of shock-responsive social protection combined remote data collection, including mobile phone surveys, with programme administrative data to generate timely evidence when in-person fieldwork was not possible.

33. Artificial intelligence enabled a rapid scale-up of efforts to consolidate and synthesize large volumes of evaluation evidence across UNICEF. The expanded use of artificial intelligence and hybrid methods also required strengthened safeguards, including human oversight, triangulation and targeted quality assurance. The Evaluation Office updated the Evaluation Evidence Gap Map (2018–2025), which systematically consolidates evaluation evidence across UNICEF programmes and identifies knowledge gaps, guiding future evaluations and reducing duplication. Offices in Afghanistan, Rwanda and Uruguay piloted artificial intelligence-supported evidence-mining and microsynthesis techniques to extract lessons from existing evaluations, while portfolio-level syntheses in Kenya, Zimbabwe, the Caribbean, Honduras and El Salvador informed programme reviews and country programme evaluations.

34. These advances extended to the function's methodological contributions to the wider United Nations evaluation community. Using text-mining techniques, it

mapped the methodologies used across more than 600 United Nations evaluation reports. The exercise generated a system-wide evidence base on evaluation methods, providing a public good for the United Nations evaluation community and supporting ongoing efforts to strengthen methodological transparency and quality across the system. Partnerships also enabled methodological innovation, including climate-resilient WASH evaluations in Eastern and Southern Africa and South Asia that applied geospatial analysis, satellite imagery and climate risk modelling through collaboration with environmental research institutions.

35. Other creative advances focused on the equally important qualitative aspects of evaluation: including children's voices and agency in evaluations and communicating better about evaluations. Evaluations in Belize, Bhutan, Maldives, Nepal and Somalia strengthened the integration of children's and community perspectives in evaluation design, data collection and validation. Efforts also extended to knowledge translation: in Barbados, Mali and Venezuela (Bolivarian Republic of), UNICEF developed interactive monitoring, evaluation and learning dashboards to improve access to evaluation findings and support their use in programme reviews and partner dialogue.

D. Moving the needle externally: Catalysing a broader movement of evidence-based accountability for children

36. Lasting impact-level results at scale require collective action that goes beyond what any single organization can achieve alone. Over the 2022–2025 quadrennium, the evaluation function sought to leverage UNICEF unique comparative advantages – its mandate, global footprint and trusted partnerships with Governments – to help catalyse a broader movement of evidence-based accountability for children. Three complementary pillars drove this agenda: impact evaluation, national evaluation capacity development and global evidence synthesis. Each saw continued momentum in 2025.

Impact evaluation

37. Impact evaluation is a highly robust form of evaluation that helps UNICEF and partners understand which interventions produce measurable results for children and therefore merit continuation, adaptation and scale-up. As such, it is a vital strategic instrument for focusing finite resources to maximize return on investment, drive impact at scale, and mitigate risk. UNICEF completed five impact evaluations during the year, maintaining steady analytical output despite fiscal constraints and in ways that added value to national policy dialogue. In Uganda, the *Impact and Economic Evaluation of Health System Strengthening Interventions in the West Nile Sub-region* generated evidence on the combined effects of cash transfers and system strengthening interventions in a fragile context, informing resource allocation decisions. In Malawi, the *Social Cash Transfer Programme Impact Evaluation (First Midline)* provided interim findings that guided programme adjustments. In Peru, the *Impact Evaluation of the High Secondary Transfer (TAS) of the Joint Programme* assessed effects on secondary education outcomes, contributing to national policy dialogue.

38. To incentivize greater investment in impact evaluation, UNICEF established the Impact Catalyst Fund as an instrument for supporting impact evaluations in priority programme areas. The fund helps country offices generate evidence to inform policy and scale decisions by financing feasibility work, baseline data collection and methodological preparation. It supported thematic windows on child marriage, social norms and adaptive social protection, with baseline data collected in Burkina Faso, Ghana, Nigeria, Somalia and Türkiye. Preparations for impact evaluations also advanced in Pakistan, Uganda and Yemen, alongside multi-country feasibility

assessments in areas such as mental health and psychosocial support, climate-resilient WASH and the PROSPECTS partnership for migrant and displaced populations, to assess readiness and evaluability before committing to full impact evaluation designs. Technical assistance requests from 14 countries reflected sustained demand for rigorous evidence despite resource constraints.

39. Impact evaluation also deepened collaboration across the United Nations system. As co-chair of the United Nations Evaluation Group working group on impact evaluation, UNICEF co-led the development of a common definition and shared methodological guidance for impact evaluation across United Nations agencies – a public good for the system that will strengthen the rigour and comparability of impact evidence for children going forward.

National evaluation capacity development

40. Whereas impact evaluation generates evidence on what works, national evaluation capacity development builds the institutional capacity of Governments and other partners to demand, produce and use their own evaluation evidence. This distinction matters: sustainable accountability for children ultimately depends on national systems, not external ones. UNICEF leverages its mandate, global reach and trusted government relationships to help embed evaluation in national planning and budgeting processes, shifting from stand-alone training towards systemic change.

41. At the global level, UNICEF continued to strengthen national evaluation systems through the Executive Course for Evaluation Leaders (ExCEL), which builds evaluation capacity among senior policy leaders (e.g. deputy ministers, parliamentarians, offices of the president and other policymakers). Two new cohorts were delivered in 2025, bringing the total to 12 cohorts and 408 certified participants from 83 countries. In several countries, ExCEL participants have gone on to support country-led evaluations, national evaluation strategies and stronger use of evaluation in government decision-making. A key milestone was Brazil's first national adaptation of the course, delivered in Portuguese with the National School of Public Administration of Brazil and the National University of Singapore. The programme also convened 39 senior officials through a subregional delivery in the Democratic Republic of the Congo. Growing demand and repeat participation suggest ExCEL is building a broader movement for evidence-based accountability for children among Governments.

42. The Intermediate Moderated Programme for Evaluation Systems' Strengthening (IMPrESS) supports government officials and national evaluation practitioners towards strengthening the institutions and practices that enable evaluation to be used in public decision-making. The programme combines structured training with practical application to help participants integrate evaluation into government planning and accountability systems. Since inception, IMPrESS has delivered 41 cohorts to 1,695 participants across 101 countries; in 2025 alone, 13 cohorts reached 694 participants. Preparatory work advanced national institutionalization, including adaptation within the École Nationale d'Administration in Benin, scheduled for launch in 2026.

43. These programmes are complemented by country-level engagement to translate new skills and networks into institutional reforms. In Mongolia, UNICEF supported collaboration with development partners and national evaluation associations to strengthen institutionalization of evaluation across government structures. In Thailand, technical assistance informed development of a national education sector evaluation framework. UNICEF also supported national evaluation policy development and system strengthening in Bangladesh, Egypt, Morocco and South

Africa. In Benin, UNICEF supported the seventh National Evaluation Day, reinforcing dialogue among government, civil society and evaluation networks.

Global evaluation synthesis

44. Global synthesis increasingly complements individual evaluations by consolidating lessons to identify recurring patterns in performance, common challenges and areas where approaches are delivering results for children. By bringing together findings from multiple evaluations, synthesis provides decision makers with a broader perspective on recurring themes, gaps and opportunities for improvement. Country offices in Afghanistan, Kenya, Lebanon, Rwanda, Uruguay and Zimbabwe conducted AI-supported rapid evidence reviews of existing evaluations to consolidate lessons across programme areas. These exercises translated accumulated findings into decision-ready insights aligned with planning and budgeting cycles, while reducing reliance on new stand-alone studies.

45. UNICEF co-chaired the Global SDG Synthesis Coalition with the United Nations Development Programme, convening 45 independent evaluation offices across the United Nations system and multilateral development banks. Under this platform, the People Pillar synthesis on social assistance interventions for Sustainable Development Goals 1–5 consolidated evidence from 155 peer-reviewed studies and 94 United Nations evaluations, and presented policy-relevant insights at high-level forums, including the Second World Summit for Social Development in November 2025. Together, these three pillars – generating rigorous evidence on what works, building the national systems to produce and use it, and synthesizing accumulated knowledge for global policy audiences – reflect a coherent and ambitious strategy for making evaluation count for children at scale.

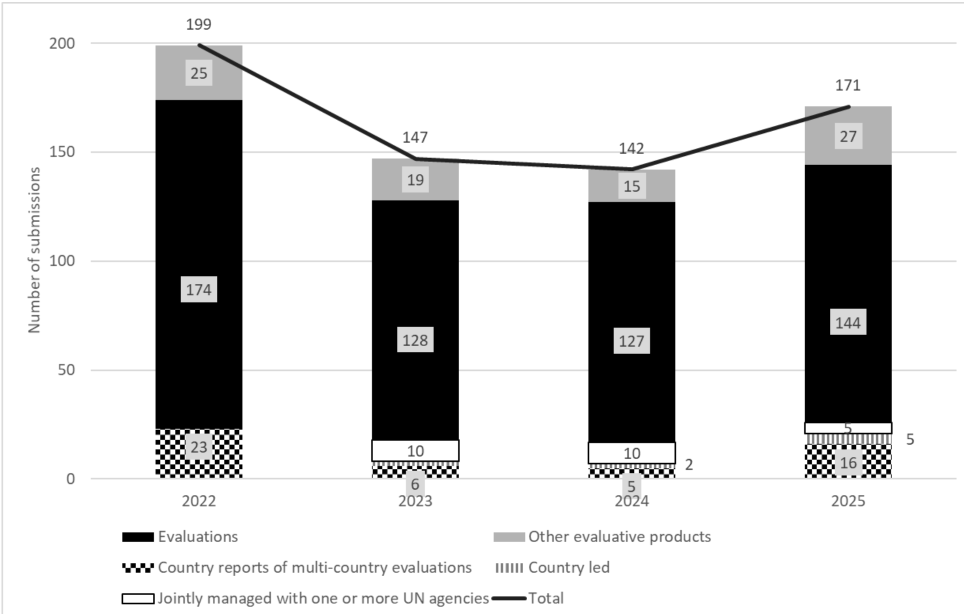
III. Performance overview: Ongoing progress and key lessons for the evaluation function

46. Throughout the quadrennium, UNICEF monitored and reported progress against the key performance indicators established for the evaluation function. The overall picture is one of broad continuity under pressure: neither the advances of earlier years nor the disruptions of 2025 dramatically shifted the trajectory of the function's performance. In a year of unprecedented organizational change, sustaining that continuity was itself a meaningful achievement. The indicators bear this out across submissions, quality, management responses and expenditure.

A. Number of evaluation submissions and geographical coverage

47. As emphasized in previous reports, the objective is not to increase evaluation numbers as an end in itself, but to ensure that evaluations are strategically relevant, credible and useful while country offices meet minimum coverage expectations. Evaluation submissions increased in 2025, following two consecutive years of managed decline. A total of 171 evaluations and other evaluative products were submitted in 2025, up from 142 in 2024.

Figure I
Number of evaluation submissions, 2022–2025



Source: Evidence Information Systems Integration (EISI) database.

48. With the high number of submissions in 2022 representing an exceptional peak driven by earlier efforts to expand coverage and a surge in COVID-related reviews, the declines in 2023 and 2024 reflect deliberate efforts to enforce clearer standards for what constitutes evaluation and a strategic shift towards greater emphasis on quality and relevance over volume. It also reflects localized resource constraints and low investment in evaluation.

49. In 2025, restructuring, staffing transitions and tighter budgets placed further deflationary pressure on evaluation activity. The observed increase therefore reflects countervailing structural and managerial factors. Standard programme cycle requirements played an important role: country programme evaluations, which are critical to informing programme design and positioning ahead of new programme cycles, more than doubled, from nine in 2024 to 20 in 2025. Multi-country evaluations also increased, reflecting efforts to examine cross-country programme performance and generate lessons applicable across regions.

50. Earlier efforts to address coverage gaps also contributed to the increase. A total of 39 country offices that reported no evaluation submissions in 2024 submitted one or more evaluations in 2025, indicating progress in extending coverage to previously underrepresented offices. At the same time, evaluability assessments increased, signalling greater use of upstream diagnostics to sequence evaluation work and ensure programme readiness before commissioning larger evaluations.

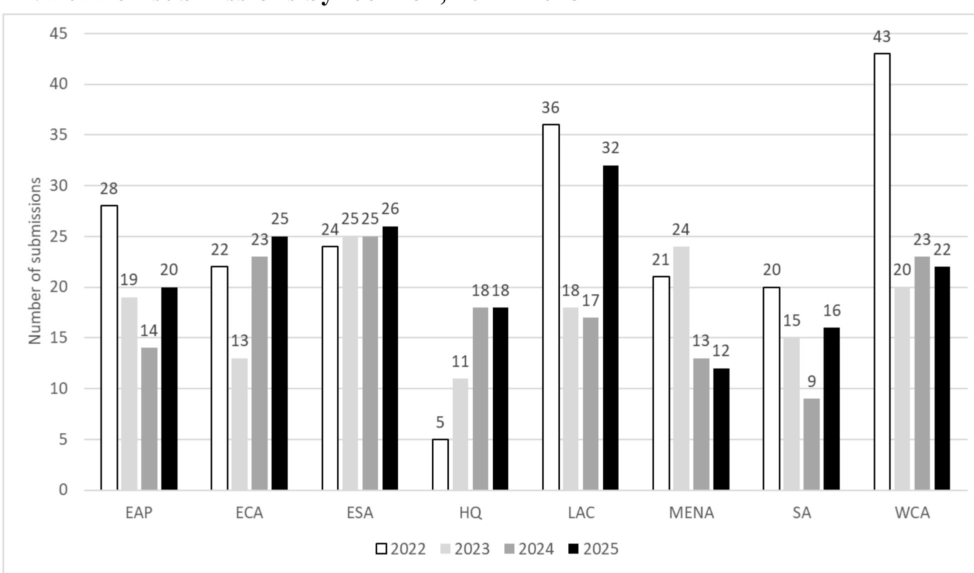
Table 1
Number of evaluations and other evaluation products submitted, by location, 2024–2025

Location	Products submitted																					
	Evaluations		Of which, country led	Of which, UNSDCF/UN system-wide	Of which, jointly managed with one or more UN agencies	Of which, country reports of multi-country evaluations	Evaluability assessments		Real-time assessments		Synthesis		Impact feasibility		Baseline assessment		Review		Total	
	2024	2025	2025	2025	2025	2025	2024	2025	2024	2025	2024	2025	2024	2025	2024	2025	2024	2025	2024	2025
EAP	10	17	3	0	0	1	3	1	0	0	0	1	0	0	1	1	0	0	14	20
ECA	23	23	1	0	1	6	0	1	0	0	0	0	0	0	0	0	0	1	23	25
ESA	25	24	0	0	1	2	0	0	0	0	0	0	0	0	0	2	0	0	25	26
HQ	14	15	0	3	1	1	0	0	0	0	2	0	1	2	0	1	1	0	18	18
LAC	16	20	0	0	0	2	1	11	0	1	0	0	0	0	0	0	0	0	17	32
MENA	11	10	1	0	1	0	2	1	0	0	0	1	0	0	0	0	0	0	13	12
SA	7	13	0	0	0	2	0	2	0	0	2	0	0	0	0	1	0	0	9	16
WCA	21	22	0	0	1	2	1	0	0	0	1	0	0	0	0	0	0	0	23	22
Total	127	144	5	3	5	16	7	16	0	1	5	2	1	2	1	5	1	1	142	171

Legend: EAP: East Asia and Pacific; ECA: Europe and Central Asia; ESA: Eastern and Southern Africa; HQ: headquarters; LAC: Latin America and Caribbean; MENA: Middle East and North Africa; SA: South Asia; WCA: West and Central Africa.
Source: EISI database.

51. Submission patterns across regions varied over the quadrennium, reflecting differences in evaluation cycles, coverage levels and operating contexts. At the same time, 13 country offices reported no submissions in both 2024 and 2025, indicating that coverage gaps remain. Regional variations largely reflect differences in evaluation cycles and operating conditions rather than a shift towards higher evaluation volumes. The pace of recovery was also influenced by staffing continuity, commissioning timelines and internal coordination. Where planning remained stable, volumes rebounded more quickly; where restructuring or staffing transitions continued, recovery was slower.

Figure II
Evaluation submissions by location, 2022–2025

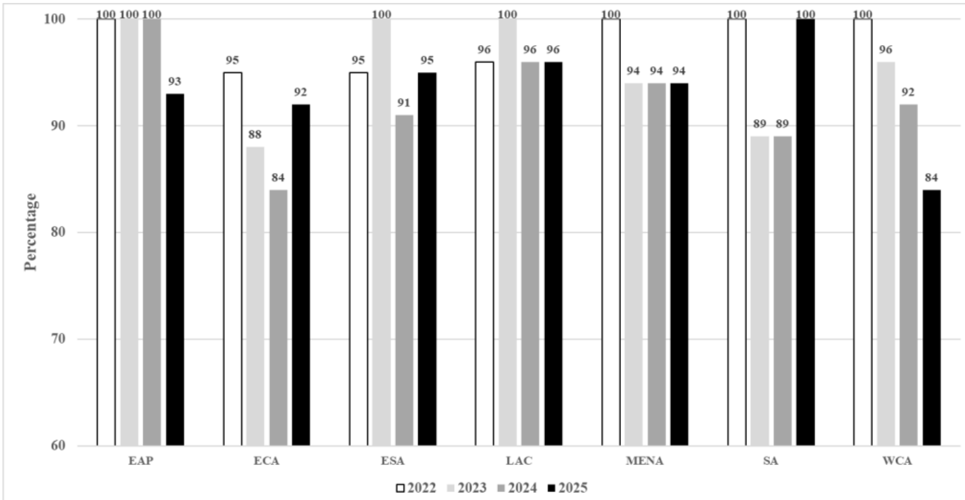


Legend: EAP: East Asia and the Pacific; ECA: Europe and Central Asia; ESA: Eastern and Southern Africa; HQ: headquarters; LAC: Latin America and Caribbean; MENA: Middle East and North Africa; SA: South Asia; WCA: West and Central Africa.
Source: EISI database.

52. Evaluation coverage – the share of country offices submitting at least one evaluation within three years – remained high across most regions in 2025. South Asia maintained full coverage, while Latin America and the Caribbean, Eastern and Southern Africa, the Middle East and North Africa, East Asia and the Pacific (93 per

cent), and Europe and Central Asia achieved nearly full coverage, while West and Central Africa maintained high, though somewhat lower, coverage than other regions.

Figure III
Evaluation coverage rates by region, 2022–2025



Legend: EAP: East Asia and the Pacific; ECA: Europe and Central Asia; ESA: Eastern and Southern Africa; HQ: headquarters; LAC: Latin America and Caribbean; MENA: Middle East and Norh Africa; SA: South Asia; WCA: West and Central Africa.
Source: EISI database.

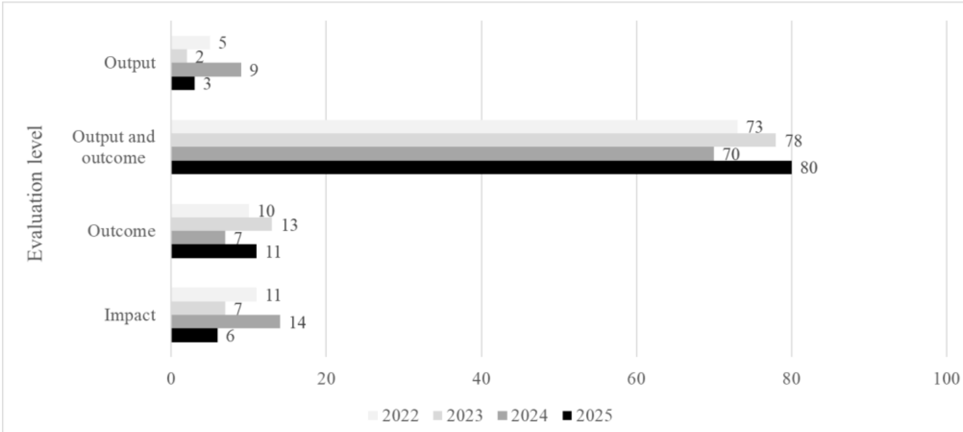
B. Evaluation profile

53. The evaluation policy commits UNICEF to a balanced evaluation portfolio, one whose evaluations increasingly assess the organization’s outcome- and impact-level results, maintain a summative and formative approach, achieve broad thematic coverage across programmatic areas, and seize on opportunities for joint, inter-agency, system-wide and country-led evaluation wherever appropriate and feasible.

Balance of purpose and thematic focus

54. In 2025, a record high number of evaluations took an outcome- or impact-level focus (97 per cent), continuing the upward trend seen since 2023. At the same time, the number focused specifically on impact remained limited, reflecting the technical challenges and cost implications associated with this level of analysis.

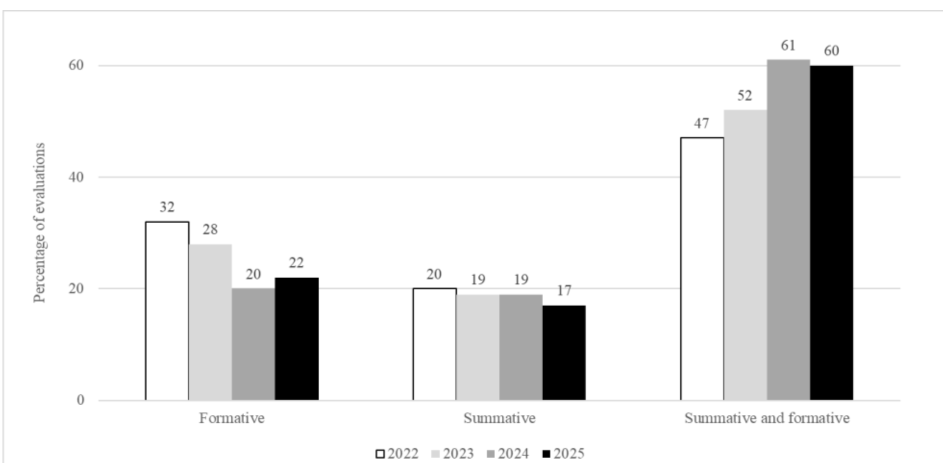
Figure IV
Evaluations by level and year, 2022–2025



Source: Global Evaluation Reports Oversight System (GEROS).

55. Formative evaluations emphasize learning for forward-looking improvement; summative evaluations emphasize accountability by assessing results accrued over relatively longer periods. Most evaluations should aim for both. The profile of evaluations' purpose shifted markedly over the quadrennium, increasingly favouring blended approaches. This trend reflects a deliberate strategic shift of the revised evaluation policy. It also reflects practical considerations: under tight budgets, integrated designs allowed offices to meet performance assessment requirements while informing programme adjustments, especially through country programme evaluations. In complex or humanitarian contexts, this blended approach is particularly crucial.

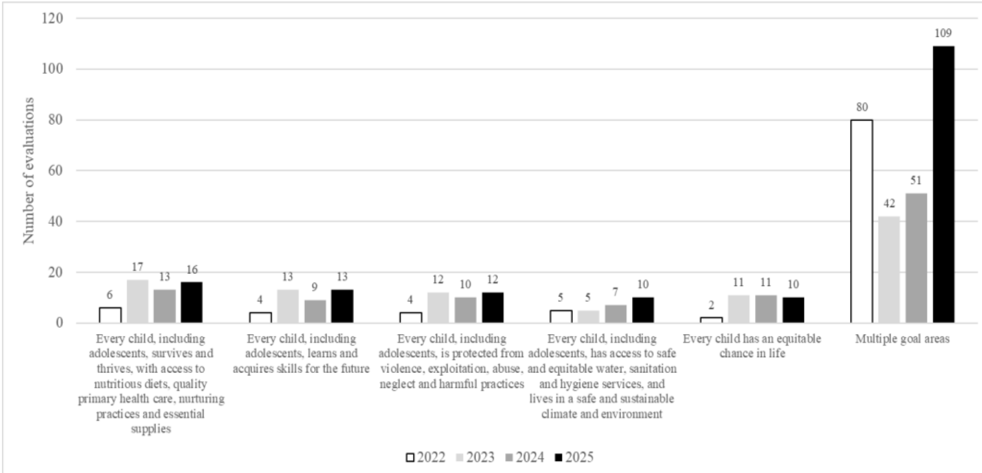
Figure V
Primary purpose of evaluations, 2022–2025



Source: GEROS.
Note: The sum does not equal 100 per cent since meta-evaluations were removed as they represent less than 1 per cent.

56. The thematic distribution of evaluations remained similarly balanced across programmatic areas, with a significant increase in evaluations covering multiple Goal Areas, which increased to their highest level in 2025. This increase reflects a deliberate push towards multisectoral approaches, in line with the organization's growing emphasis on integrated programming, an emphasis reflected in national partners' increasing demand for integrated policy insights.

Figure VI
Evaluation focus, by UNICEF Goal Areas, 2022–2025



Source: EISI database.

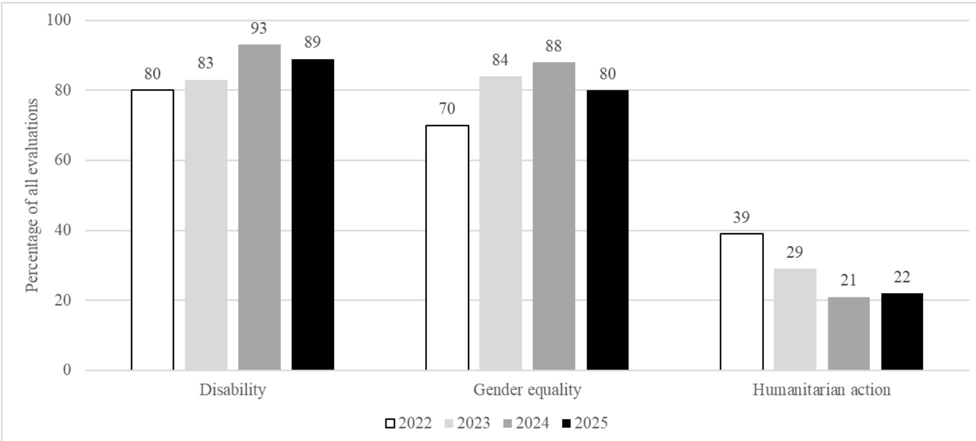
Joint, inter-agency, system-wide and country-led evaluations

57. UNICEF continued to undertake joint, system-wide and country-led evaluations in 2025, though numbers remained limited. Five evaluations were jointly managed with other United Nations agencies, three were system-wide and five were country-led. Compared to previous years, joint evaluations declined, while country-led and system-wide evaluations increased modestly. This pattern mainly reflects structural constraints, including misaligned planning cycles, differing programmatic priorities, high coordination costs, resource constraints and the decentralized nature of evaluation commissioning. Partnerships with Governments and United Nations agencies nevertheless remained strong, especially in capacity development, joint programming and evidence generation. Thus, the rise in country-led evaluations suggests growing national engagement.

Evaluation coverage of disability, gender equality and humanitarian action

58. Integration of the cross-cutting lenses of disability, gender equality and humanitarian lenses into evaluations remained high overall. This said, progress was variable. After an upward trend in the first three years of the quadrennium, integration of disability and gender equality declined slightly in 2025.

Figure VII
Evaluations covering disability, gender equality and humanitarian action, 2022–2025

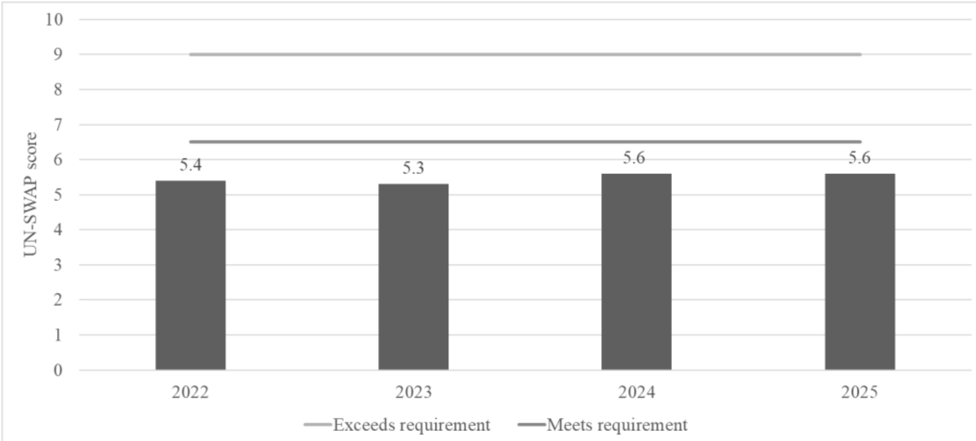


Source: GEROS.

59. The humanitarian lens is categorically different than the other two cross-cutting themes. Its integration into evaluations has historically been low, and declining since its peak in 2022, for several reasons. First, the 2022 was an anomaly, owing to the number of COVID-19-related evaluative exercises completed that year. Second, some countries do not lend themselves as readily as others to the application of this lens because their profile is heavily development-focused (in which case a systematic integration of humanitarian issues is somewhat less relevant). Others are intrinsically humanitarian in profile, and are therefore either covered by evaluations that are entirely humanitarian in scope (in which case the humanitarian lens defines the evaluation in the first instance), or are such large-scale complex emergencies that they undertake ongoing reviews to support operational adjustments and risk management, rather than retrospective performance reviews. Third, the meaningful integration of the humanitarian lens requires specialized skills, so even in some cases where it is called for, doing so is constrained by limited staff capacity, uneven guidance and coordination challenges with humanitarian partners.

60. The United Nations System-wide Action Plan on Gender Equality and Women's Empowerment index measures the meaningful integration of gender in evaluations. Despite sustained efforts in 2024 and 2025 to address persistent underperformance on this criterion, performance remained at its 2024 level, consistently approaching but not reaching the “meets requirement” threshold. Constraints include gaps in sex- and age-disaggregated data, monitoring systems not designed for gender analysis, and uneven access to specialized expertise. Gender was often referenced in design or treated as a stand-alone section, but not consistently integrated across evaluation questions, methods, findings and recommendations. Reliance on generalist teams and compressed timelines further limits systematic gender-responsive approaches.

Figure VIII
Evaluation performance under the United Nations System-wide Action Plan on Gender Equality and Women's Empowerment, 2022–2025



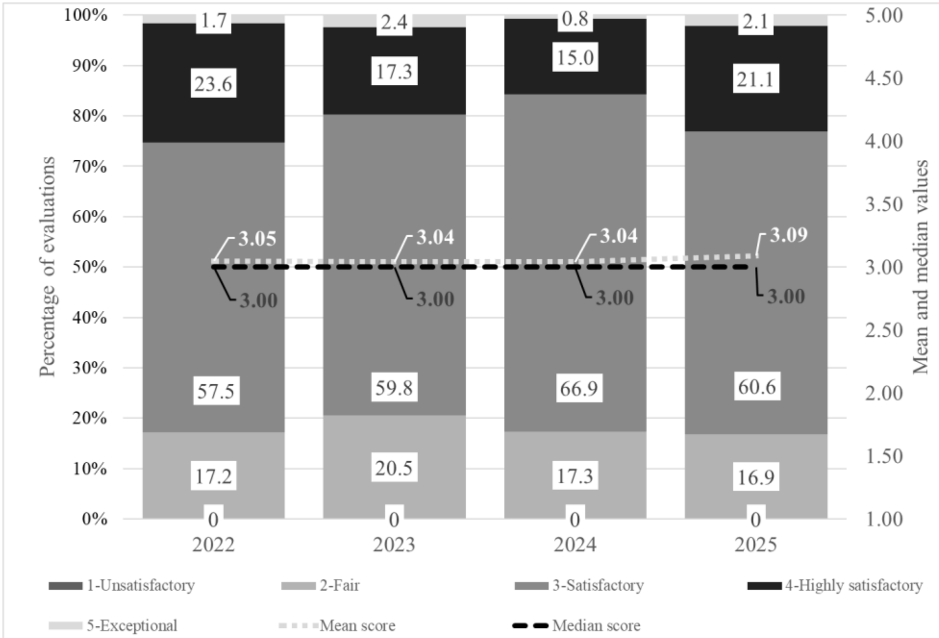
Source: GEROS.

C. Quality of evaluations

61. Evaluation quality remained stable across the 2022–2025 quadrennium and improved slightly in 2025. The mean quality score increased from 3.04 in 2024 to 3.09 in 2025, while the median score remained stable at 3.0 across the quadrennium, confirming overall consistency as shown in figure IX.

62. Most evaluations met satisfactory quality standards, while a growing share reached "highly satisfactory" or "exceptional", indicating solid overall quality with scope to further strengthen analytical depth, methodological rigour and the clarity of recommendations. Differences in quality were closely linked to enabling conditions: compressed timelines, limited management capacity, and fewer opportunities for iterative technical review tended to reduce analytical depth, while strong methodological and strategic guidance at the inception stage, clear definition of stakeholder roles, and early leadership and technical support improved quality.

Figure IX
Overall evaluation quality ratings, 2022–2025



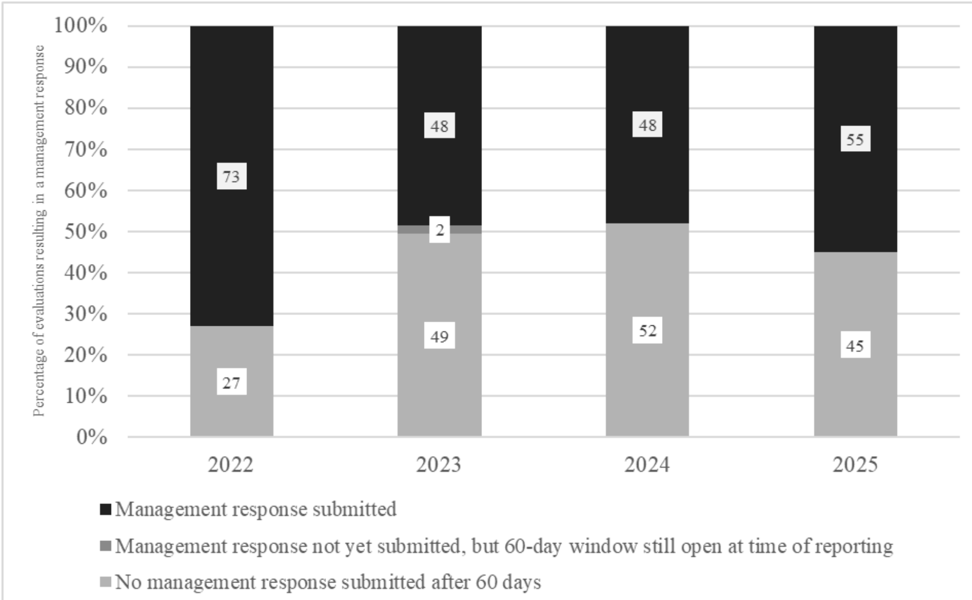
Source: GEROS.

D. Evaluation management response submission and implementation

63. In accordance with the evaluation policy, management responses, which convey the extent to which management accepts each recommendation and outline how recommendations will be addressed, are expected within 60 days of the issuance of a final evaluation report. Although the timely delivery of a management response is a helpful proxy for evaluation use, it does not necessarily mean that an evaluation's recommendations have been meaningfully acted on. Conversely, the lack of a timely management response does not always mean that its recommendations are not being meaningfully acted on; on the contrary, some of the most meaningful actions can sometimes occur before a management response is delivered.

64. Response patterns during the year suggest some improvement compared with previous years, though a significant share of evaluations still exceeded the 60-day window or remained pending at the time of reporting. These patterns should also be viewed in the context of FFI-related organizational disruptions, which affected staffing continuity and coordination. Overall, the data suggest gradual strengthening of response discipline, while highlighting the need for more consistent monitoring and stronger systems to track meaningful evaluation use.

Figure X
Timeliness of management response submissions, 2022–2025

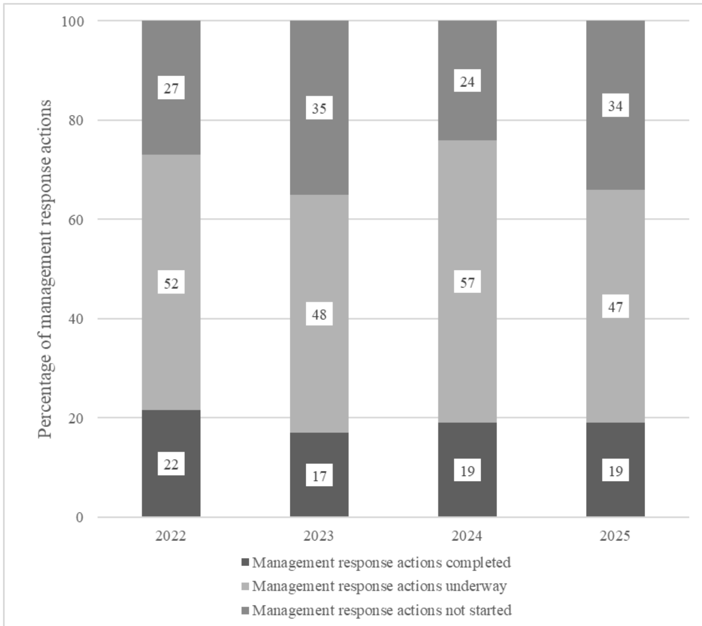


Source: EISI database.

65. Tracking implementation of management responses provides insights into whether agreed actions are translating into the changes prescribed in evaluation recommendations. Implementation patterns remained broadly stable throughout the quadrennium, with many recommendations completed, while others remained in progress or not yet started, reflecting the time required to translate evaluation findings into operational and policy adjustments.

66. Overall, while the system for acting on evaluation recommendations is functioning, stronger prioritization, monitoring and accountability mechanisms are needed to sustain momentum.

Figure XI
Management response actions and implementation status, 2022–2025

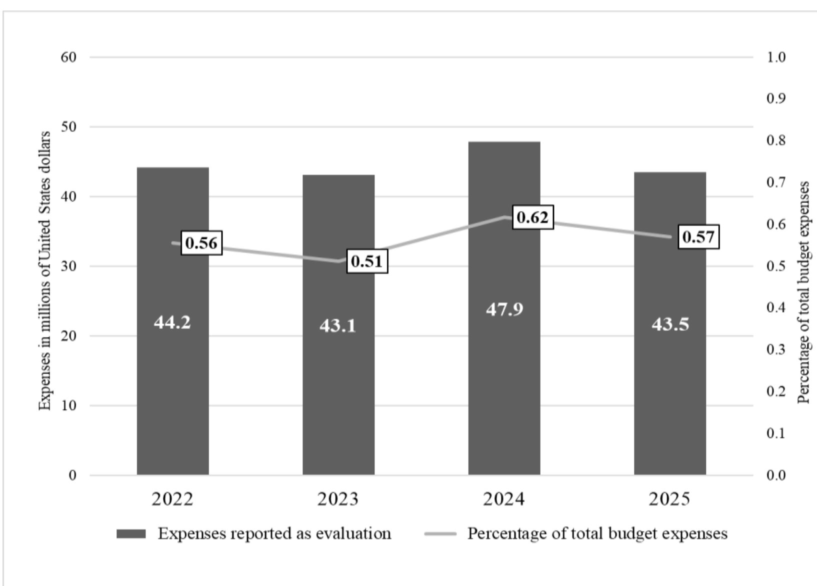


Source: EISI database.

E. Evaluation expenditure

67. As a policy commitment, UNICEF aims to allocate 1 per cent of programme expenditure to evaluation. During the quadrennium, expenditure remained broadly stable but below this benchmark.

Figure XII
UNICEF expenditure reported as evaluation, in millions of United States dollars and as a percentage of total expenses, 2022–2025

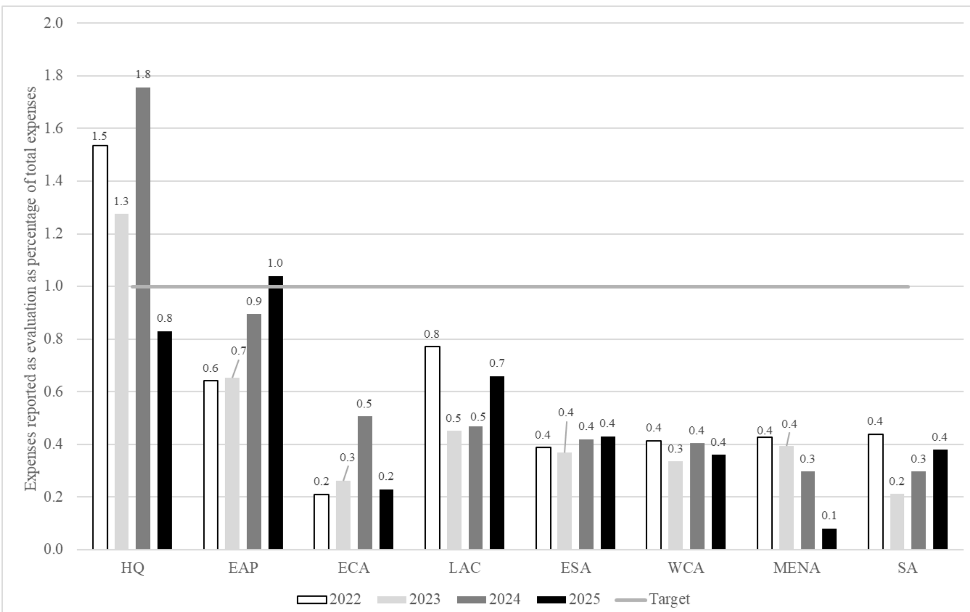


Note: The bars show evaluation expenditure reported each year (US$ millions). The line shows evaluation expenditure as a percentage of total expenses, with values labelled for each year.
Numerator: evaluation expenditure; denominator: total net programme expenditure.
Source: VISION extract (6 March 2026).

68. Reported expenditure does not directly track the number of evaluations completed. Although more evaluations were reported in 2025, changes in evaluation expenditure did not move in a simple one-to-one relationship with the number of evaluations delivered. This reflects differences in evaluation scale, scope, timing, complexity, portfolio management and cost-coding practices across offices and organizational levels.

69. Patterns remained variable across the function, with some locations periodically reaching the corporate benchmark and most regions remaining below it. Year-to-year volatility suggests that changes reflect portfolio timing, cost structures and reporting practices rather than shifts in evaluation activity alone. In two cases (the Evaluation Office and Europe and Central Asia), spikes observed in 2024 were not sustained in 2025 despite increased submissions, underscoring the notion, described above, that higher volumes were delivered within similar or tighter resource envelopes.

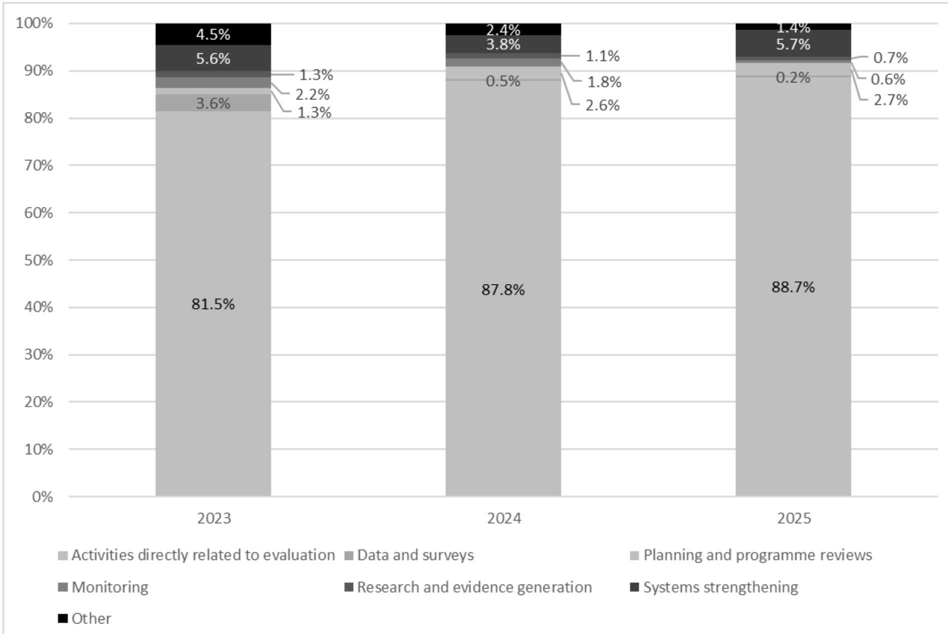
Figure XIII
Expenditure reported as evaluation as a percentage of total expenses, by region, 2022–2025



Legend: East Asia and the Pacific (EAP); Europe and Central Asia (ECA); Eastern and Southern Africa (ESA); Headquarters (HQ); Latin America and Caribbean (LAC); Middle East and North Africa (MENA); South Asia (SA); West and Central Africa (WCA).
Source: EISI database.

70. Year-to-year changes partly reflect how expenditures are coded, and inconsistent application of guidance limits comparability of reported investment across the organization. Trends in expenditure coding suggest gradual improvement in the accuracy of how evaluation costs are recorded. Over the quadrennium, most categories previously susceptible to a high degree of miscoding saw increased accuracy, while the share of expenditure correctly coded as evaluation-related correspondingly increased. This indicates that deliberate guidance and monitoring efforts to improve coding practices have begun to take effect. One notable exception is the "systems strengthening" category, which remained relatively large and rebounded after an earlier decline. This category may partly reflect investment in national evaluation capacity development activities; further analysis will be undertaken to clarify its composition and ensure that coding practices accurately capture evaluation-related investments.

Figure XIV
Activities reported as evaluation expenditure, by activity type, 2023 to 2025



Source: Evaluation expenditure data retrieved from VISION (6 March 2026).

71. Taken together, the performance indicators for the 2022–2025 quadrennium tell a coherent story: a function that made genuine progress in the areas most directly within its control – evaluation quality, outcome-level focus, blended methodological approaches and coding accuracy – while continuing to face structural constraints in areas that depend on broader organizational enabling conditions, including management response timeliness, evaluation expenditure levels and the consistent integration of cross-cutting lenses. The resilience of these patterns in the face of the significant pressures of 2025 suggests that the institutional foundations laid over the quadrennium are holding. At the same time, the persistent gaps point clearly to the priorities that will require continued attention as the function enters the new quadrennium.

IV. Conclusions and way forward

72. The 2022–2025 quadrennium began at a difficult juncture for the evaluation function: in the immediate aftermath of a year-long leadership gap and an extended period of remote operations during the COVID-19 pandemic, and only shortly after the decentralized level of the function had been established within UNICEF. Under new leadership, the quadrennium was defined by an ambitious effort to reassemble and reinvigorate a dispersed team, unify the three levels of a recently globalized function under a shared vision, and develop and implement a revised evaluation policy that would reposition the function for optimal impact. This ensures that evaluations at all levels are relevant, timely, credible and useful; that the function operates with the independence that underpins its integrity; and that adequate, predictable resources are provided to enable it to fulfil its mandate.

73. As momentum was building, the onset of FFI introduced an unprecedented period of institutional change and financial uncertainty. FFI required the function to adapt and reprioritize – participating actively in the change process, contributing evaluative evidence to support the organization's navigation of a difficult transition,

and simultaneously sustaining progress on policy implementation, delivering on the commitments of the Plan for global evaluations, 2022–2025, and developing a successor plan calibrated to the new organizational realities.

74. The results documented in this report reflect what the function achieved against that demanding backdrop. On the quantitative indicators, the story is one of resilience: performance remained broadly stable across submissions, quality, management responses and expenditure, with no significant backsliding despite the pressures of 2025 – a result that was neither inevitable nor easily won. On the qualitative dimensions, the story is one of genuine progress. The function's core work continued with a sharpened emphasis on utilization, producing evaluations that demonstrably influenced institutional reform, programme design, national policy and strategic decision – including, for the first time in such direct and substantive terms, the development of the UNICEF Strategic Plan, 2026–2029. The hard-wiring and soft-wiring of a maturing function advanced significantly: the first-ever *UNICEF Evaluation Handbook* was completed, GELT crystallized as a vehicle for coherence and collective stewardship; and strategic initiatives in impact evaluation, national evaluation capacity development and global synthesis continued to build the broader movement of evidence-based accountability for children that is the function's most ambitious long-term objective. And the Plan for global evaluations, 2026–2029 – smaller, better-prioritized and more transparent than any of its predecessors – provides a strong foundation for the quadrennium ahead.

75. Building on these foundations, the evaluation function enters the new quadrennium with a clear set of priorities, namely:

(a) **Ensuring implementation of the highest-priority commitments** in the Plan for global evaluations, 2026–2029 while attempting to mobilize resources for the remaining priorities;

(b) **Sustaining focus on outcome- and impact-level results measurement** in evaluations, and seeking creative methodological approaches to doing so ever more robustly;

(c) **Attaining consistently high relevance, timeliness, credibility and utility of evaluations** by socializing the completed *UNICEF Evaluation Handbook*; expanding the suite of technical guidance tools to build advanced methodological skills; developing templates; strengthening quality assurance processes, beginning at the evaluation design stage; and better-customized internal capacity development efforts, spearheaded by a newly formed Global Evaluation Learning Committee and grounded in a systematic skills-mapping across the function;

(d) **Forging a stronger utility-to-use linkage, together with management**, through improved timeliness and monitoring of management responses, stronger accountability for implementation, more systematic tracking and documentation of meaningful evaluation use, targeted follow-up on select evaluations, and continued use of evaluation synthesis to elevate recurring issues to management's attention;

(e) **Reinforcing the significant gains made in 2025 in moving the needle for children beyond UNICEF**, and doing so in an increasingly constrained financial environment, by continuing its leadership in national evaluation capacity development, impact evaluation and evaluation synthesis, and responding to Governments' increasing demands for country-led evaluations;

(f) **Strengthening transparency in evaluation financing** through costed multi-year planning, improved expenditure tracking and reporting, and risk-based prioritization of resources across regions and country offices; and

(g) **Building on the significant internal cohesion, coherence and collaboration forged though the creation of GELT**, which will work together to harmonize the global evaluation practice, and will serve as a platform for aligning strategic approaches and navigating ongoing changes and developments within the organization and in its operating environment.

76. The function will also continue to monitor structural changes within UNICEF, including the establishment of Centres of Excellence, the creation of the Office of Strategy and Evidence, and the consolidation of regional offices. Ongoing discussions on governance reform and internal oversight arrangements may further influence the institutional context in which the evaluation function operates. As these changes reshape the broader evidence ecosystem and the institutional arrangements through which evaluation engages with programme leadership, supports decentralized evaluation and contributes to organizational learning, the function will continue to adapt as it did in 2025. Importantly, an evaluation of organizational transformation, enshrined in the Plan for global evaluations, 2026–2029, will include an assessment of the FFI as one of several change-management initiatives undertaken by the organization. With an internal after-action review commissioned by the Office of the Executive Director also taking place in 2026, the Evaluation Office will continue to work closely with the after-action review team to minimize duplication and maximize complementarity between the two exercises.

77. One particularly significant opportunity on the horizon is the forthcoming independent external assessment of evaluation in UNICEF, to be conducted by OECD-DAC in response to Executive Board decisions 2023/12 and 2023/20. Assessing the health and maturity of the function and its enabling environment from a fully independent and impartial perspective, the assessment will help strengthen evaluation practice and inform the review and revision of the evaluation policy in 2027–2028. OECD-DAC will engage broadly with UNICEF personnel, the Executive Board and the Audit Advisory Committee throughout, and will present its findings to the Executive Board at its first regular session of 2027. For a function that has spent the better part of a quadrennium building its foundations under difficult conditions, this assessment represents both a moment of accountability and an opportunity – a chance to take stock of how far the function has come, identify where further strengthening is needed, and chart a course for evaluation that continues to deliver for children in the years ahead.

V. Draft decision

The Executive Board

Takes note of the annual report for 2025 on the evaluation function in UNICEF (E/ICEF/2026/17) and its management response (E/ICEF/2026/18).