



# Economic and Social Council

Distr.: General  
14 April 2025

Original: English

**For decision**

---

## United Nations Children's Fund

Executive Board

**Annual session 2025**

10–13 June 2025

Item 12 of the provisional agenda\*

### **Annual report for 2024 on the evaluation function in UNICEF**

#### *Summary*

This report provides an overview of the UNICEF evaluation function in 2024. 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 decision for consideration by the Executive Board are provided in section V.

---

\* E/ICEF/2025/11.



## I. Overview

1. This year witnessed positive momentum overall, both within the evaluation function and in its enabling environment. The Revised evaluation policy of UNICEF ([E/ICEF/2023/27](#) and [E/ICEF/2023/27/Corr.1](#)) went into effect on 1 January 2024, signalling a set of strategic shifts to further strengthen the evaluation function so that, through its dual learning and accountability role, it is optimally positioned to help UNICEF and its partners accelerate progress for children in the critical years ahead. During this first year of the revised policy, the evaluation function and management worked together to begin actioning both the “hard-wiring” and “soft-wiring” elements that will help ensure consistently robust implementation of the policy throughout the organization.
2. In parallel with this initial year of change, the evaluation function continued its core work of generating timely, relevant, credible and impactful evidence to inform the strategic direction of the organization throughout the year. Its contributions to key decision-making processes culminated in the Evaluation of the UNICEF Strategic Plan, 2022–2025 ([E/ICEF/2025/3](#)), which was presented to the Executive Board at its first regular session of 2025, and which the organization has used extensively in developing the UNICEF Strategic Plan, 2026–2029.
3. Progress on the key performance indicators established for evaluation continued, with several noteworthy positive advances being observed. The volume of evaluations stabilized from the prior year, owing in part to concerted efforts to reduce this number by prioritizing the most strategically relevant themes. Declines at the regional and country levels were offset by a significant increase in the number of evaluations at headquarters level, especially in the areas of institutional effectiveness and humanitarian action.
4. Overall quality remained satisfactory to highly satisfactory, notwithstanding variations across the function. However, while evaluations more systematically incorporated gender and disability lenses, the meaningful incorporation of gender concerns remained unacceptably low despite an intensive effort to improve performance in this area.
5. One particularly encouraging development was an increase in evaluation expenditure, both in absolute terms and as a percentage of total programme expenses. Expenditures at headquarters exceeded the 1 per cent organizational target, owing to an advance of resources from the Evaluation Pooled Fund that had originally been allocated for 2025 but were drawn down to cover 2024 evaluation priorities. Moreover, all but one region recorded an increase in expenditure from 2023, with one region (East Asia and the Pacific) approaching the 1 per cent target. This reversal of the prior year’s pattern of decline was the direct result of strong collaboration among the evaluation function (at both the global and decentralized levels), regional directors and many country representatives. This progress will need to be sustained, however; in most regions, evaluation expenditures still did not reach their pre-decline levels, and in some countries, expenditures remained unacceptably low in proportion to their size and other aspects of their risk profile.
6. Alongside these efforts within the organization itself, UNICEF continued to catalyse a broader global movement of learning and accountability to accelerate progress for children. Its national evaluation capacity development (NECD) efforts scaled up significantly: through its innovative flagship course for senior policymakers, by the end of 2024 UNICEF had equipped 224 high-level government officials from 70 countries with the knowledge and skills to engage substantively in evaluation in ways that will facilitate better-informed decision-making processes for the benefit of children. UNICEF scaled up efforts to embed impact evaluation as a

tool for making strategically sound decisions on where and how to achieve the highest return on investment for children. Together with the World Food Programme (WFP), UNICEF hosted the Global Impact Evaluation Forum 2024 at the United Nations, which has catalysed strategic partnerships among organizations from the inter-governmental, non-governmental, private, public and philanthropic sectors. Finally, UNICEF played a leadership role in fostering global partnerships on evaluation synthesis, a strategic priority in an increasingly complex decision-making environment. Co-chairing the Global SDG Synthesis Coalition, UNICEF played a central role in the governance of the Evidence Synthesis Infrastructure Collaborative by mobilizing \$74 million for evaluation synthesis.

7. The harnessing of innovative and creative data-driven solutions to maximize evaluations' timeliness, strategic relevance, credibility, utility, use, efficiency and cost-efficiency was another hallmark of progress in 2024. These solutions ranged from process improvements (e.g. data-driven forensic analysis to understand the root causes of performance shortfalls, development of a risk-based evaluation planning tool, engagement of evaluation advisory groups comprised of young people) to the streamlining of data systems (e.g. increased use of administrative data systems, user-focused improvements to the organization's evaluation database), methodological advances (e.g. modelling tools to forecast future performance scenarios) and the use of new technologies (e.g. artificial intelligence and natural language processing). An initiative was also launched to enhance the function's strategic communications capabilities to maximize evaluation impact.

8. Amid these significant steps forward, several challenges persisted in 2024. In addition to the ongoing variability of performance across the function on key indicators, the challenge of measuring organizational achievements beyond outputs and activities continued, owing to a lack of robust outcome- and impact-level data. This gap extended to the evaluation function itself: beyond anecdotal accounts, there was little robust evidence on the scope or scale of the broader impact of evaluations on the work of the organization. The most significant shortfall was the delay in some strategic initiatives: with finite capacity to meet the manifold challenges and opportunities the function faces, the development of three key documents that complement the revised evaluation policy – i.e. the evaluation policy procedure document, the policy implementation strategy and the evaluation handbook – were started in 2024 but their completion was deferred to 2025.

9. In 2025 and beyond, efforts to address persistent challenges while sustaining the gains of 2024 will continue. These efforts will take place within a context of evolving organizational realities for UNICEF. Within this period of change, the overarching focus of the evaluation function will be twofold: (a) to support the organization, through relevant evaluative evidence, in navigating a complex and uncertain landscape; and (b) to maintain the integrity of evaluation as a robust independent function, with the appropriate structures and adequate and predictable resources that enable it to fulfil its dual learning and accountability role, so that it remains future-fit amidst the wider currents reshaping the organization.

## **II. Key developments in the evaluation function**

### **A. Background and priorities for 2024**

10. The years prior to 2024 had been a period of flux and transition for the evaluation function. After the issuance of a revised evaluation policy in 2018 and the function's corresponding expansion from 2018 to 2019 through the establishment of the decentralized function (aided by the creation of the Evaluation Pooled Fund), the COVID-19 pandemic manifested itself. During this period the Director of Evaluation

position became vacant, leaving the newly established, geographically dispersed and remotely operating global team without leadership for nearly a year.

11. It was against this backdrop that in late 2021, with the arrival of a Director of Evaluation, the key priority was to (re)build the team, technically and relationally, and to establish consistency of good practice for an independent evaluation function in accordance with the United Nations Evaluation Group (UNEG) norms and standards. Alongside this priority was another, higher-order strategic priority: a policy-mandated review and revision of the evaluation policy in 2023, preceded by independent peer review of the function in 2022.

12. The priority in the two years prior to 2024 was therefore to undertake a systematic, inclusive, evidence-based approach to concluding these processes and developing the Revised evaluation policy of UNICEF ([E/ICEF/2023/27](#) and [E/ICEF/2023/27/Corr.1](#)), which was approved by the Executive Board at its second regular session of 2023 and came into effect on 1 January 2024. The consistent, robust and meaningful implementation of the revised policy was therefore the overarching priority of 2024 – while continuing to build an integrated and harmonized global evaluation function, ensure uniformly high performance, strengthen the enabling environment and evaluation culture, and implement an ambitious evaluation work plan, all within an organizational structure in which staff of the independent function at decentralized level do not report to the Director of Evaluation.

13. The revised policy embodies a series of incremental but significant changes to mature the evaluation function and its enabling environment and thus optimize it to help UNICEF and its partners achieve the best possible results for children in the years ahead. It further demarcates specific strategic priorities to future-fit the function for maximum impact for children. These include NECD, impact evaluation, evaluation synthesis, strategic evaluation partnerships and strategic communications.

14. The broad strategy for ensuring consistently robust implementation of the policy within the organization's decentralized structure followed two parallel tracks, roughly categorized as the “hard-wiring” and “soft-wiring” aspects of action. The “hard-wiring” elements consisted of the physical outputs that operationalize the broad goals of the policy in more concrete and granular detail. These include three key documents: (a) the policy implementation strategy, which provides the overarching road map for fulfilling the commitments set forth in the policy; (b) the policy procedure document, which prescribes the actions and expectations assigned to specific individuals responsible for implementation; and (c) the evaluation handbook, which provides evaluation staff and others engaging in evaluation detailed technical “how-to” for producing consistently relevant, timely, credible and useful evaluations. All three of these documents were initiated in 2024, but their finalization was deferred to 2025 owing to capacity limitations.

15. The “soft-wiring” element, by contrast, focused on forging and deepening the relationships most crucial for internalizing the policy – and importantly, the principles that underpin it – and ensuring its meaningful implementation. This element included relationship-building across and within the three levels of the organization. Owing to particularly close collaboration between the evaluation function (Evaluation Office and regional evaluation advisers) and management (regional directors, division directors, and the Office of the Executive Director), considerable progress was made in this area. Roll-out of the policy was included in the agenda of most regional management team meetings, and trilateral meetings were established between regional directors (or deputy regional directors), regional evaluation advisers and the Director of Evaluation to simultaneously strengthen the independence of the function, its close embeddedness to management and decision-making processes, and its overall performance at the decentralized level.

## B. Noteworthy evaluations contributing to results

16. Amid these overarching strategic currents, UNICEF continued to deliver evaluations that contributed to the organization's performance in delivering results for children.

17. The Evaluation of the UNICEF Strategic Plan, 2022–2025 (E/ICEF/2025/3) was a particularly noteworthy global evaluation, both for its leveraging of creative solutions to ensure a timely yet thorough and robust assessment of the entire quadrennium and for its positive impact on organizational decision-making processes. The creative solutions it relied on centred on methodological approaches that not only enabled the delivery of a high-quality evaluation in just five months, but rather also enabled the evaluation team to forecast outcomes likely to be observed by the end of 2025. Accordingly, the organization has used the results of this evaluation extensively and thoughtfully – including, most prominently, in the development of the draft UNICEF Strategic Plan, 2026–2029.

18. Other, more narrowly focused global evaluations of the organization's broad strategic approaches proved similarly impactful. The *Evaluation of the Implementation of the UNICEF Gender Policy and Gender Action Plan (GAP 2 and GAP 3)* identified areas for gender equality mainstreaming across the humanitarian, development and peace contexts, and informed the next Gender Action Plan and its integration into the next Strategic Plan. The evaluation of the reasonable accommodation fund found that the fund effectively fostered a more inclusive work environment by removing workplace barriers for employees with disabilities, while also recommending a range of improvements.<sup>1</sup> The *Developmental Evaluation of UNICEF's Internal Communication and Staff Engagement Function*, meanwhile, recommended a shift beyond one-way communication towards a more dynamic, collaborative approach that promotes engagement and belonging. The *Evaluation of UNICEF's Work in Contributing to Child Poverty Reduction, Including in Humanitarian and Fragile Settings* proved similarly insightful. In all of these evaluations, UNICEF management was actively, positively and constructively engaged, and has been acting on their recommendations.

19. The number of evaluations of humanitarian action increased in tandem with the long-standing growth in the organization's role in emergencies across the globe. Altogether, the Evaluation Office undertook 12 evaluations of emergency responses – more than double the number produced in 2023 – which generated strategically and operationally relevant evidence to inform the organization's work. The *Evaluation of UNICEF's Level 3 (L3) Response in Ukraine* highlighted good practices such as shifting to a targeted, needs-based implementation model, as well as the integration of short-term aid with long-term recovery efforts by leveraging local capacity. Similarly, the *Evaluation of the UNICEF L3 Response in Yemen* noted progress from 2019 to 2022, but recommended improvements to humanitarian tools, enhanced programme coherence, a reassessment of emergency procedures, leveraging of field experience, and strengthened oversight and partnerships to better target vulnerable groups and ensure sustainable interventions. The Evaluation Office played a key role in developing the first *UNEG Guidance on the Integration of Humanitarian Principles in the Evaluation of Humanitarian Action*.

20. A wide range of evaluations, too numerous to summarize, were produced at the decentralized level, aiding in decision-making processes within the organization's operations on the ground. Noteworthy examples include the *Evaluation of UNICEF's Response to the Level 2 Cholera and Cyclone Freddy Emergencies in Malawi*, which recommended regionalizing and strengthening supply chain management and

<sup>1</sup> All evaluations in this report can be accessed at <https://www.unicef.org/evaluation/reports/>.

procurement strategies to build local resilience and reinforce humanitarian-development-peace linkages in future emergencies. The *Summative Evaluation of the EU Global Promotion of Best Practices for Children in Migration Programme* raised the profile of the topic and explored the implications of the evaluation for future programming. *Inter-Agency Humanitarian Evaluation of the Response to the Crisis in Northern Ethiopia and Afghanistan* assessed the added value of the whole-of-United-Nations response, thus informing the decision-making of the humanitarian coordination teams in both countries. The *Evaluation of the UNICEF Pacific Child Protection Programme 2018–2023: Kiribati, Samoa, the Solomon Islands and Vanuatu* led to a significant grant extension in mid-2024.

21. Beyond these individual evaluations, UNICEF built its evaluation synthesis and meta-analysis practice in 2024 as a means of elevating to stakeholders' attention the issues that repeatedly arise in evaluations. The *Global Evaluation Evidence Synthesis: Evidence of UNICEF Achievements for Children from 2018–2022 Evaluations* analysed the outcomes of UNICEF work from 2018 to 2022, leveraging artificial intelligence (AI) to map evaluation results by goal areas, change strategies, enablers and other elements of the UNICEF Strategic Plan, 2022–2025. Its identification of critical evidence gaps will directly inform future evaluation planning processes. At a programmatic level, the *Meta-synthesis and Meta-analysis of Evidence on Child Marriage in South Asia* examined 65 studies on child marriage in South Asia – including studies of interventions beyond UNICEF itself – and identified effective strategies for reducing child marriage in the region. This analysis has since informed global discussions on evidence-based programming and interventions effectiveness beyond South Asia.

22. Although these well-evidenced positive outcomes are noteworthy, 2024 saw a continuation of a long-standing knowledge gap surrounding the overall scope and scale of UNICEF evaluation's effects as a whole. Beyond individual examples of high-quality, high-use evaluations such as these, an overall picture of the function's magnitude of influence on the work of the organization remained unclear owing to a lack of monitoring systems and capacity.

### **C. Impact evaluation**

23. Impact evaluation provides the robust evidence that Governments, donors and implementing partners need to make the most effective – and most cost-effective – investments to maximize results for children. It facilitates better decisions on whether an intervention (or combination of interventions) should be continued and scaled up, either “as is” or in modified form, or discontinued. Crucially, it helps decision makers understand which intervention (or combination of interventions) will yield the highest return on investment. As UNICEF and many other organizations strive to deliver more and better in an increasingly resource-constrained environment, impact evaluation is more critical than ever. This is why this modality is a centrepiece of the revised evaluation policy.

24. More evaluations based on impact methodology (5) and evaluations assessing impact level results (18) were completed in 2024 than in previous years. An *Impact Evaluation of Targeted Social Assistance (TSA) in Georgia* was found to have demonstrated the superior benefits of cash disbursements compared to food stamps, leading to a 30 per cent expansion in programme coverage. Similarly, the impact evaluation of the Multi-Sector Nutrition Plan in Nepal provided robust evidence to inform scale-up decisions, covering 352,000 children and fostering continued collaboration among diverse stakeholders. The positive findings of the impact evaluation of the Transfer to Early Childhood programme in Peru resulted in an

additional allocation by the Government of \$8.6 million to expand programme coverage to reach more children.

25. The Impact Catalyst Fund (ICF), a flagship initiative of the Evaluation Office, proved pivotal in incentivizing the use of impact evaluation at country level. Burkina Faso, Ghana, India and Türkiye received ICF matching grants to design and conduct rigorous impact evaluations focused on child marriage and social norms. In total, 10 country offices are now slated to conduct impact evaluations under two thematic windows: child marriage, and social norms and adaptive social protection.

26. UNICEF thought leadership to advance impact evaluation for the benefit of children culminated in the Global Impact Evaluation Forum 2024 at the United Nations, co-hosted by UNICEF and WFP in New York in December. The Forum brought together 80 organizations from 30 countries, including Member State delegations, bilateral and multilateral donors, senior officials from 12 United Nations agencies, non-governmental organizations, the private sector and philanthropies. The UNICEF Executive Director and Deputy Executive Director for Partnerships participated in the Forum, emphasizing the moral imperative to accelerate joint efforts on impact evaluation. The event resulted in actionable takeaways to guide the future of impact evaluation and catalysed a range of partnerships among participants.

#### **D. National evaluation capacity development**

27. Whereas impact evaluation can help achieve results at scale by ensuring well-targeted, evidence-driven investments in high-yield interventions, national evaluation capacity development (NECD) can help achieve systems-level results at scale by bringing other key sectors and actors beyond UNICEF into the sphere of evaluation. As UNICEF is just one among many contributors to results for children, it is only through these other stakeholders' engagement in evaluation – especially senior policymakers, as Governments are the primary duty bearers for children's rights – that learning and accountability will have far-reaching and tangible benefits for children at scale. It is for this reason that NECD, like impact evaluation, is identified as a key priority in the revised evaluation policy.

28. As with impact evaluation, UNICEF made significant strides in NECD in 2024. Delivery of the online course Intermediate Moderated Programme for Evaluation Systems Strengthening (IMPrESS) continued, with 1,187 government officials, partners, United Nations staff and other technical evaluation practitioners and managers representing 67 countries receiving training in 2024. UNICEF also scaled up the Executive Course for Evaluation Leaders (ExCEL), an innovative programme for senior policymakers delivered in partnership with the National University of Singapore that focuses on helping these officials and their staff apply an evaluation lens to find solutions for children in their countries. By the end of 2024 – just over two years since its first cohort was trained – ExCEL had welcomed 101 delegations from 70 countries representing all regions served by UNICEF. In total, 224 senior policymakers and other government partners (e.g. directors-general, deputy ministers, senior advisers) graduated from ExCEL, together with 107 UNICEF staff who accompanied them in the course. Together, these 331 ExCEL alumni constitute a network of evaluation champions who continue to foster national, intercountry, regional and global exchanges of experience.

29. Moreover, ExCEL has catalysed a broader movement of intergovernmental collaboration on evaluation. For example, the Department of Planning, Monitoring and Evaluation of South Africa leveraged the ExCEL network to invite fellow Governments to join its newly established National Evaluation System Resource Centre, which was itself created as a result of ExCEL. Following participation in ExCEL, the Government of Bangladesh sent a delegation to exchange experience with

the Government of Sri Lanka, resulting in the Enhancing Evaluation Capacity for Development bilateral agreement with UNICEF. In Turkmenistan, participation in ExCEL led to deep-dive workshops and the integration of evaluation capacity building into the National Action Plan for the Realization of Children's Rights 2023–2028, which was approved by the President. In Cambodia, UNICEF leveraged the knowledge acquired by several ExCEL alumni to develop a professional certification course for evaluation practitioners and managers through the National University of Management.

30. UNICEF also continued to meet more individualized NECD demands at country level to strengthen evaluation systems and the role of evaluation in decision-making. Examples include: (a) National Evaluation Weeks in Benin, Cameroon and Togo; (b) tailored support to the Government of Sierra Leone; (c) a National Capacity Diagnosis in Congo; and (d) a capacity-building programme in Burundi to upskill the Monitoring and Evaluation Ministerial Technical Group to conduct rapid evaluations, as well as the establishment of an evaluation community of practice hosted by the University of Burundi. New initiatives in 2024 included: (a) a Young and Emerging Evaluator learning programme in Yemen; (b) a workshop with the Office of the Civil Service Commission in Thailand to develop a curriculum for future civil servants; and (c) a training session with the Ministry of Planning and Investment in Viet Nam. UNICEF also facilitated a range of regional learning webinars and participated in numerous global and regional conferences, including the Asia Pacific Evaluation Association Conference in the Philippines, the EvalPartners Global Evaluation Forum in Morocco, the National Evaluation Capacities Conference in China, and the Academic Symposium on M&E Education in Sri Lanka.

## **E. Innovations and other creative solutions**

31. The revised evaluation policy identifies innovative and creative problem-solving as a means of maximizing the relevance, timeliness, quality, credibility, utility and efficiency of evaluations. Numerous advances were made in 2024 in the areas of evaluation planning, management and the participation of young people.

32. In preparation for the UNICEF Plan for global evaluations, 2026–2029, UNICEF began developing a risk-based planning tool that will bring greater consistency, rigour, objectivity and transparency to the identification of themes to be evaluated in the next quadrennium. This tool, which will be piloted in 2025, integrates a broad range of data and information sources and will form the basis for discussions within the function and for consultations with internal and external stakeholders. Beyond streamlining the planning process, this tool will help ensure prioritization of the most strategically relevant evaluation topics, and by extension will rationalize resource allocations. It will also facilitate a more conjoined approach to evaluation planning across the three levels of the function, while also enhancing independence by elevating the role of systematic, data-driven, evidence-based analysis in the planning process.

33. As AI-driven automation and emerging technologies continued to evolve in 2024, UNICEF advanced the use of AI to improve its evaluation processes and conduct. Improvements to data extraction, processing, analysis and sharing through the use of AI boosted efficiency and accuracy, as well as the ability to produce stronger evidence syntheses from diverse data sets, thus aiding real-time decision-making. Moreover, AI-powered enhancements to the UNICEF Evidence Information Systems Integration (EISI) search engine led to improved searchability of evaluation reports and a surge in user searches.

34. UNICEF also elevated AI as an overall strategic evaluation priority, both within the organization and beyond. Advances included the publication of “An operational



framework for machine learning in evaluation”, as well as two technical papers and a journal article in *Social Science & Medicine*. As co-chair of the UNEG Data and AI Working Group, UNICEF shared AI use cases and data science applications potentially relevant to other organizations. The Working Group finalized a discussion paper and recommendations – as well as a technical paper and recommendations prepared with the Advisory Group on Ethics on AI in Evidence Generation Activities – on ethical issues related to AI in evidence generation.

35. These “high-tech” solutions were complemented by an array of equally important “low-tech” advances. To make findings, conclusions and recommendations from evaluations in emergencies more widely accessible, UNICEF upgraded its use of visual aids to disseminate key conclusions and lessons, and tested creative alternatives for engaging stakeholders in the evaluation process, thereby promoting their sense of buy-in. For example, a “client-journey” methodology was employed during select humanitarian evaluations to better portray the lived experience of affected populations before, during and after their interactions with organizations responding to the emergency. In addition, UNICEF took a more systematic approach to elevate the voices of those it serves – particularly young people, through the establishment of a Youth Advisory Board which played an active role in the global evaluations of the Strategic Plan and the Gender Action Plan. For its exceptional contribution to meaningfully engaging young people in evaluation, UNICEF was recognized as a “super champion” among the United Nations agencies by the Eval4Action campaign, which promotes evaluation as critical to, and a key accelerator for, achieving the Sustainable Development Goals.

36. Numerous creative solutions were explored at the decentralized level as well. Teams in Peru and Ecuador tested the use of administrative data as a cost-effective approach to increase the timeliness of impact evaluations. In the South Asia region, a combination of AI-driven analysis and manual review was used to conduct a comprehensive appraisal of all UNICEF evaluations in the region over the past three years, which contributed evidence-informed analysis to the Regional Office Management Plan. Similarly, in the East Asia and Pacific region, machine learning was deployed in the form of a “co-pilot” tool that was trained to extract key words and phrases from evaluation reports to help inform an evaluation synthesis. In the Europe and Central Asia region, an adolescent-trained AI model was used to assess if government mental health services met shifting client demands. The Eastern and Southern Africa region used AI tools to produce two micro-syntheses on child protection in humanitarian action: one contributing to the evidence base for the UNICEF Strategic Programme Framework for Strengthening Civil Registration and Vital Statistics in Eastern and Southern Africa; and the other facilitating the rapid extraction of key findings from multiple evaluations to support a synthesis report on UNICEF emergency response efforts.

## **F. Partnerships**

37. Along with innovative and creative problem-solving, partnerships constitute a key strategic pillar for fulfilling the ambitious commitments set forth in the revised evaluation policy – and, by extension, for harnessing evaluation as an accelerator of progress for children. The Global Impact Evaluation Forum 2024 at the United Nations, which UNICEF co-hosted with WFP and which catalysed partnerships between and among participants, represents a touchstone achievement. Moreover, the ability of UNICEF to build its impact evaluation capacity and practice is the direct result of a creative partnership with the German Federal Ministry for Economic Cooperation and Development (BMZ) on the Adaptive Social Protection: Evidence for Child Outcomes (ASPECT) project.

38. Partnerships also feature prominently in the NECD efforts of UNICEF, particularly in design and delivery as well as the relationships these efforts have spawned among participants. The touchstone activities at the global level – ExCEL and IMPrESS – were complemented by regionally customized initiatives. The West and Central Africa Regional Office, for example, formalized a multi-year partnership with the Centre Africain d'Études Supérieures en Gestion to embed evaluation competencies into public institutions in the region, equipping Governments with advanced skills and systems to conduct and utilize evaluations for policy and programme improvements. In Sri Lanka, UNICEF established a partnership with WFP and the German Institute for Development Evaluation to conduct national diagnostics on evaluation capacity to generate a comprehensive assessment of the country's evaluation ecosystem and capacity to support evidence-based decision-making. Together with United Nations partners, the Eastern and Southern Africa Regional Office organized a five-week training course to build institutional knowledge and capacities of Governments and United Nations staff on evaluation frameworks, planning and national evaluation system development.

39. Significant gains observed in the area of evaluation synthesis were likewise the result of strong partnerships with other United Nations agencies and others. UNICEF, in collaboration with United Nations sister agencies,<sup>2</sup> synthesized lessons from all United Nations system evaluations related to Sustainable Development Goal 5 (“Achieve gender equality and empower all women and girls”). In addition, UNICEF and the United Nations Development Programme (UNDP) co-chaired the Global SDG Synthesis Coalition (GSDGSC) of more than 46 United Nations agencies. The organization also played a key role in the governance of the Evidence Synthesis Infrastructure Collaborative (ESIC), a new initiative funded by the Wellcome Trust to address societal challenges through expanded use of evaluative evidence, safe and responsible use of AI, and increased data sharing and reuse. Owing in large part to the success of this partnership and through the initial investments from UK Research and Innovation as well as Wellcome Trust \$74 million was mobilized in support of GSDGSC and ESIC – not for UNICEF itself but rather for the broader global endeavours.

40. Joint, inter-agency and system-wide collaborations constituted yet another area of marked progress. Ten evaluations were jointly managed with other agencies, including evaluations of the COVID-19 Vaccines Global Access initiative,<sup>3</sup> the Global Action Plan for Healthy Lives and Well-being for All,<sup>4</sup> and the joint WHO/UNICEF evaluation of the Global Task Force on Cholera Control. UNICEF provided support to the United Nations Sustainable Development Group System-Wide Evaluation Office – fully operational as of 2024 – and identified areas for coordination and collaboration. This included the production of five United Nations evaluation evidence summaries for status reporting on the Quadrennial Comprehensive Policy Review (QCPR), as well as the use of AI to generate interactive evidence maps of all United Nations evaluation reports at the global, regional and country levels since 2021. These maps of United Nations system evaluation evidence and evidence summaries provided evidence on several topics, namely: (a) the resident coordinator system; (b) the United Nations development system regional architecture; (c) sustainable food systems; (d) humanitarian-development-peace linkages; and (e) funding quality. They garnered interest and positive feedback from Member States and provided key information to stakeholders involved in the 2024 QCPR.

---

<sup>2</sup> United Nations Entity for Gender Equality and the Empowerment of Women (UN-Women), the United Nations Development Programme (UNDP), the United Nations Population Fund (UNFPA) and WFP.

<sup>3</sup> Together with Gavi, the Vaccine Alliance, and the World Health Organization (WHO).

<sup>4</sup> Together with a global coalition of 13 organizations, led by WHO.

41. UNICEF remained actively involved in global and regional evaluation networks. Engagement in UNEG working groups led the development of several guidance documents on the integration of humanitarian principles, human rights and gender equality into evaluations. Other key contributions were made in the areas of ethics in evaluation; UNEG evaluation synthesis; environmental and social impact; methods; data and artificial intelligence; gender, disability and human rights; peacebuilding; impact evaluation; decentralized evaluations and professionalisation; young and emerging evaluators; humanitarian evaluation; and system-wide United Nations Disability Inclusion Strategy working groups. In partnership with the United Nations Population Fund (UNFPA), UNICEF contributed to the synthesis report on the United Nations Youth Strategy: Priority area education and employment. The United Nations Evaluation Development Group for Asia and the Pacific, a regional inter-agency network of which UNICEF is a core member, worked closely with the United Nations Development Coordination Office to deliver a training for Governments and United Nations country teams in countries that had scheduled evaluations of United Nations Sustainable Development Cooperation Frameworks.

### **III. Performance overview: ongoing progress and key lessons for the evaluation function**

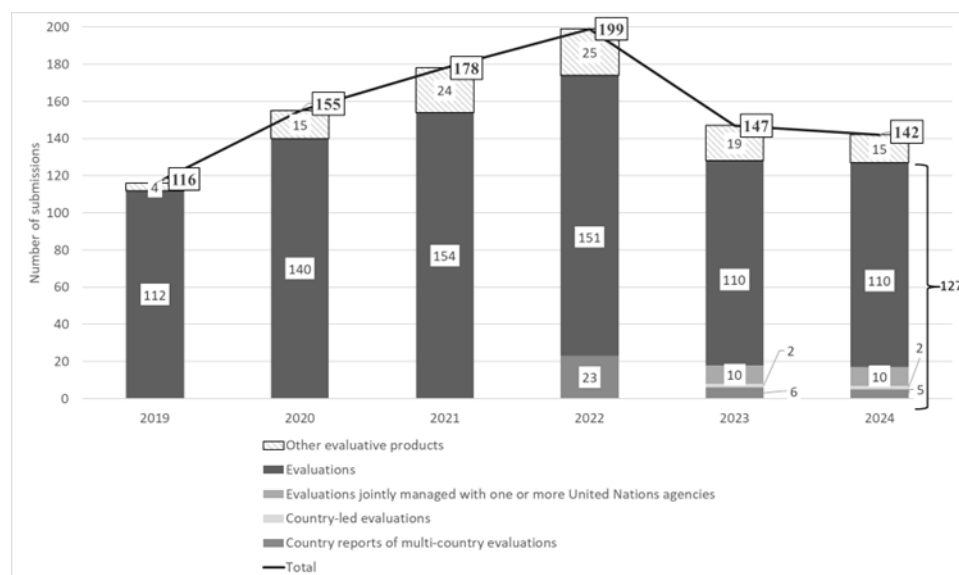
42. This section presents an overview of the major trends observed in relation to the key performance indicators for evaluation in 2024. In keeping with a commitment made in the revised evaluation policy, the UNICEF Strategic Plan, 2026–2029 will offer an opportunity to develop additional metrics that holistically capture the qualitative health and maturity of the function and its enabling environment in complement to the existing quantitative indicators.

#### **A. Number of evaluation submissions and geographical coverage**

43. As figure I indicates, 142 evaluations and other evaluative outputs were recorded in 2024. These include 127 evaluations – 10 of which were jointly managed with other United Nations agencies, 5 multi-country evaluations and 2 country-led – and 15 other evaluative products.

44. The sustained stabilization in the number of evaluations at a level comparable to pre-pandemic levels is noteworthy. This stabilization is likely the result of numerous factors, including: (a) clearer criteria for classifying an exercise as an evaluation, introduced in 2022 and supplemented in 2023 and 2024 with a verification exercise to confirm accuracy of submissions; (b) a broad strategic appeal, enshrined in the revised evaluation policy, to prioritize timeliness, strategic relevance, quality and credibility over quantity; and (c) a decline in several regions owing to resource constraints, under-resourcing of evaluation and other local decision-making processes.

Figure I  
Number of evaluation submissions, 2019–2024



Note: Breakdown of types of evaluations was introduced in 2023.

Source: EISI database.

45. Another factor contributing to declining volume was the reduction in the number of standalone evaluability assessments. Although their numbers declined, 2024 saw an uptick in the incorporation of these early, rapid and future-oriented exercises into standard evaluations, suggesting a positive trend towards more integrated, holistic and balanced evaluations that simultaneously address immediate forward-looking learning needs as well as more comprehensive, in-depth assessments of performance.

Table 1  
Number of evaluations and other evaluation products submitted, by location, 2022–2024

Location	Products submitted																
	Evaluations			Of which, country led	Of which, jointly managed with one or more United Nations agencies	Of which, country reports of multi-country evaluations	Evaluability assessments			Real-time assessments			Other (includes baseline assessment, synthesis, review)			Total	
	2022	2023	2024	2024	2024	2024	2022	2023	2024	2022	2023	2024	2022	2023	2024	2023	2024
EAP	23	13	10	1	0	0	2	5	3	3	1	0	0	0	1	19	14
ECA	16	11	23	1	2	5	3	2	0	3	0	0	0	0	0	13	23
ESA	24	24	25	0	2	0	0	0	0	0	0	0	0	1	0	25	25
HQ	4	11	14	0	0	0	1	1	0	0	0	0	0	1	4	13	18
LAC	27	17	16	0	0	0	0	1	1	9	0	0	0	0	0	18	17
MENA	21	21	11	0	1	0	0	2	2	0	0	0	0	1	0	24	13
SA	18	14	7	0	2	0	2	1	0	0	0	0	0	0	2	15	9
WCA	41	17	21	0	3	0	2	3	1	0	0	0	0	0	1	20	23
<b>Total</b>	<b>174</b>	<b>128</b>	<b>127</b>	<b>2</b>	<b>10</b>	<b>5</b>	<b>10</b>	<b>15</b>	<b>7</b>	<b>15</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>3</b>	<b>8</b>	<b>147</b>	<b>142</b>

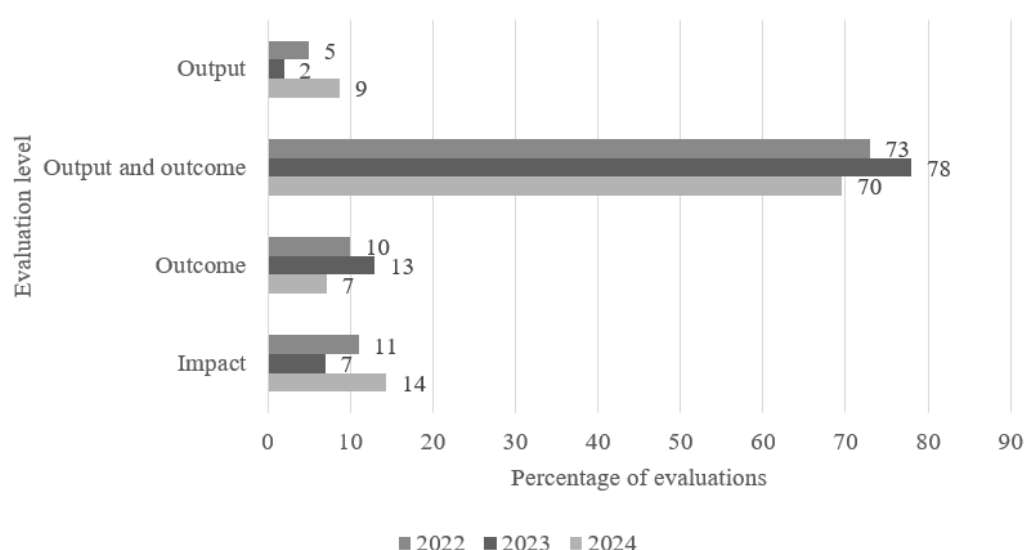
Legend: EAP: East Asia and the Pacific; ECA: Europe and Central Asia; ESA: Eastern and Southern Africa; HQ: headquarters; LAC: Latin America and Caribbean; MENA: Middle East and North Africa; SA: South Asia; WCA: West and Central Africa.

Source: EISI database.

46. Further reflecting the appeal for a focus on strategically relevant evaluations rather than a higher volume, 2024 saw a continuation of a prior trend whereby a large majority of evaluations focused on higher-order outcomes and impacts on the lives of

children. Sub-analysis beyond the data displayed in figure II furthermore indicates that, among the 11 country offices conducting output-level evaluations in 2024, 7 had also undertaken impact- and outcome-level evaluations. This pattern suggests the prevalence of a holistic and balanced blend of evaluative perspectives in these offices, rather than a default to less ambitious output-level evaluations.

Figure II  
Evaluations by level and year, 2022–2024

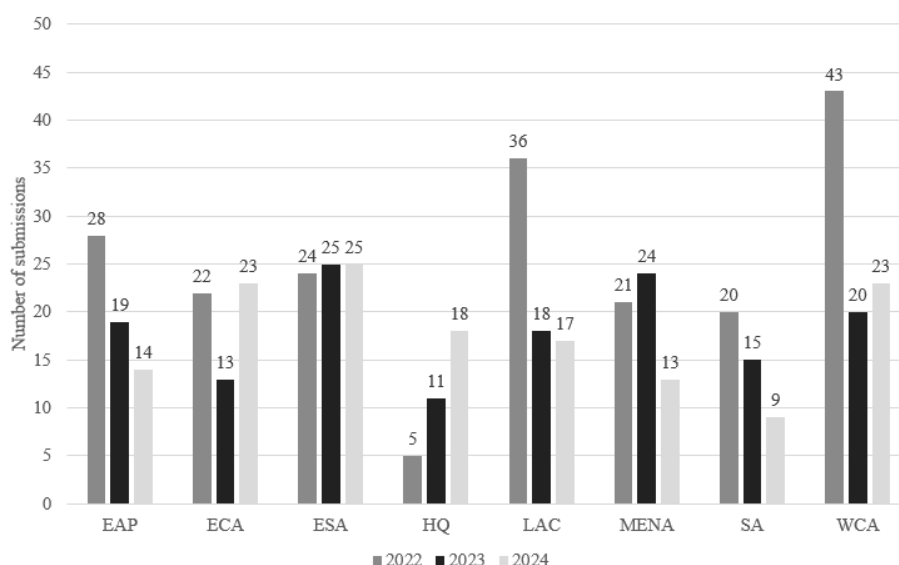


Source: Global Evaluation Reports Oversight System (GEROS).

47. Figure III summarizes trends in evaluation submissions by location. Within the broader trend of sustained reduction in volume, output varied significantly across the organization. Submissions at headquarters more than tripled, owing to the need to address gaps in evaluation coverage of strategic criticality, coupled with increasing demand for evaluation at this level of the organization. This uptick in demand is arguably a positive indication of growing appreciation and value placed on evaluation; it also a testament to the willingness of the function to accommodate requests and meet demands and, owing to the various innovative and creative measures it deployed to achieve efficiencies, its ability to deliver on time and at a consistently high level of quality. This absorptive capacity was stretched to its limit in 2024, however.

48. Output levels varied greatly from one region to the next. Europe and Central Asia and West and Central Africa surpassed their 2023 output levels. Eastern and Southern Africa submitted the most evaluations overall, while South Asia (which covers just eight countries) submitted the fewest. Significant declines occurred in Middle East and North Africa, South Asia, and East Asia and the Pacific. The fluctuations in evaluation submissions in Middle East and North Africa are linked primarily to challenging programme environments, particularly in L2 and L3 emergency contexts. In South Asia and East Asia and the Pacific, declines were offset by greater integration of country-level into global or thematic evaluations and the longer duration entailed by these regions' higher participation in impact evaluations. Notably, while volume declined in these regions, Global Evaluation Reports Oversight System (GEROS) ratings improved somewhat, illustrating the potential gains from a shift from quantity to quality.

Figure III  
Evaluation submissions by location, 2022–2024



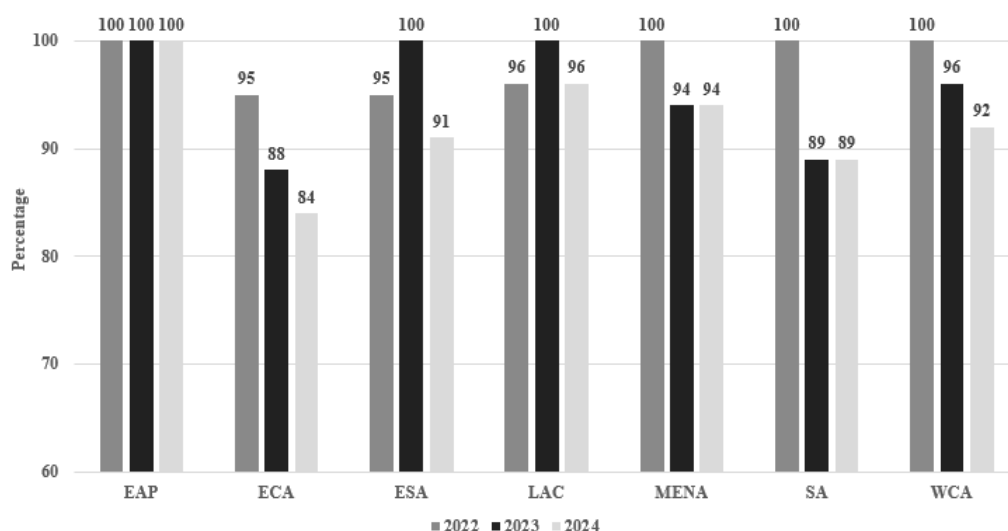
*Legend:* EAP: East Asia and the Pacific; ECA: Europe and Central Asia; ESA: Eastern and Southern Africa; HQ: headquarters; LAC: Latin America and Caribbean; MENA: Middle East and North Africa; SA: South Asia; WCA: West and Central Africa.

*Source:* EISI database.

49. Evaluation coverage, summarized in figure IV, was uneven, with a notable decline in 2024 compared to 2023 in most regions. This is due to as many as 24 countries across various regions failing to submit evaluation reports in the past three years. East Asia and the Pacific maintained full coverage despite a decline in overall submissions, as offices that had not conducted evaluations in the previous two years did so in 2024. Conversely, Europe and Central Asia saw a sharp decrease in coverage even though the number of evaluation submissions in this region increased significantly, as some offices conducted multiple evaluations. Coverage levelled off in South Asia and Middle East and North Africa, but a significant drop in coverage was observed in other regions except East Asia and the Pacific in 2024.

50. The most concerning observation in 2024 was the outright lack of evaluation activity in more than 30 countries, which together represent roughly one quarter of all UNICEF programmes. Only 5 of these 30 offices were operating in L2 or L3 countries, and most managed significant resources: 7 offices possessed budgets between \$25 million and \$50 million, 3 possessed budgets of \$51 million to \$100 million, and 3 had budgets over \$100 million. These gaps persisted despite the “soft-wiring” tactics employed by the evaluation function and management to encourage country offices to meet their obligations.

Figure IV  
Evaluation coverage rates by region, 2022–2024



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

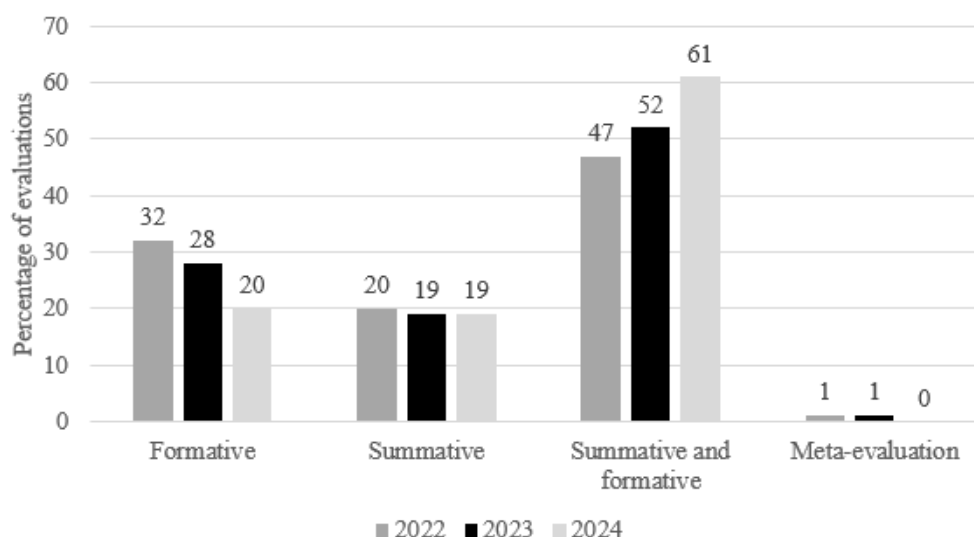
*Source:* EISI database.

## B. Evaluation profile

### 1. Balance of purpose and thematic focus

51. Some evaluations are primarily formative in nature, aiming to provide stakeholders “feed-forward” relatively early in the programme cycle, and others are primarily summative in nature, aiming to provide a more definitive assessment of performance and granular feedback at a later period in the programme cycle. A balanced blend of both forms is ideal, and many evaluations provide insights of both types. As figure V illustrates, the profile of evaluations in 2024 continued to be well balanced.

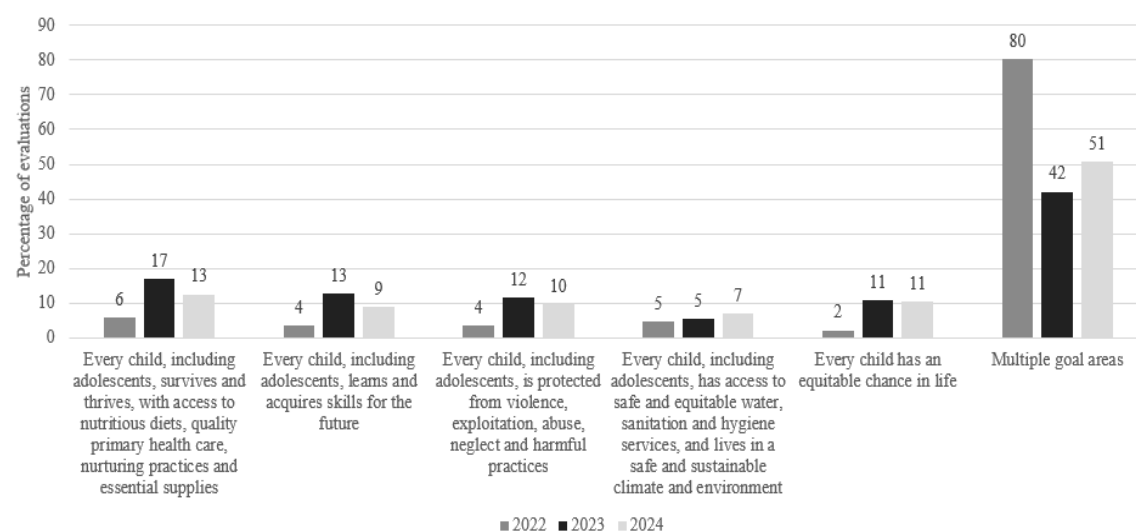
Figure V  
Primary purpose(s) of evaluations, 2022–2024



Source: GEROS.

52. Balance in coverage of the UNICEF Goal Areas is also desirable. Owing to the holistic nature of children's development – and the intersectionality of risks to children, with multiple sources of vulnerability entailing heightened risk and a corresponding need for integrated programmatic approaches – it is incumbent upon the function to ensure that evaluations examine performance in an integrated manner rather than perpetuating siloes. In 2024 modest progress was made to reverse a trend, witnessed the year prior, of declining multisectoral evaluation coverage.

Figure VI  
Evaluation focus, by UNICEF Goal Area(s), 2022–2024



Source: EISI database.

## 2. Joint, inter-agency and country-led evaluations

53. A balance is also desirable in ensuring evaluation coverage of UNICEF itself and of its work with partners. In 2024, UNICEF engagement in this latter category – through joint, inter-agency, system-wide and country-led evaluations – increased. Similar to the previous year, a total of 12 evaluations – including Inter-Agency



Humanitarian Evaluations – were jointly managed with other United Nations agencies such as WHO, UNFPA, WFP, the Office of the United Nations High Commissioner for Refugees, the Resident Coordinator Office, the International Organization for Migration, and the United Nations Office on Drugs and Crime, in various countries and covering thematic areas ranging from child protection to child marriage, girls' education and global health. In addition to its general support to the System-Wide Evaluation Office, UNICEF supported specific system-wide evaluations, including evaluations of the Spotlight Initiative, which is committed to ending violence against women and girls, and the United Nations Disability Inclusion Strategy, which is currently under way.

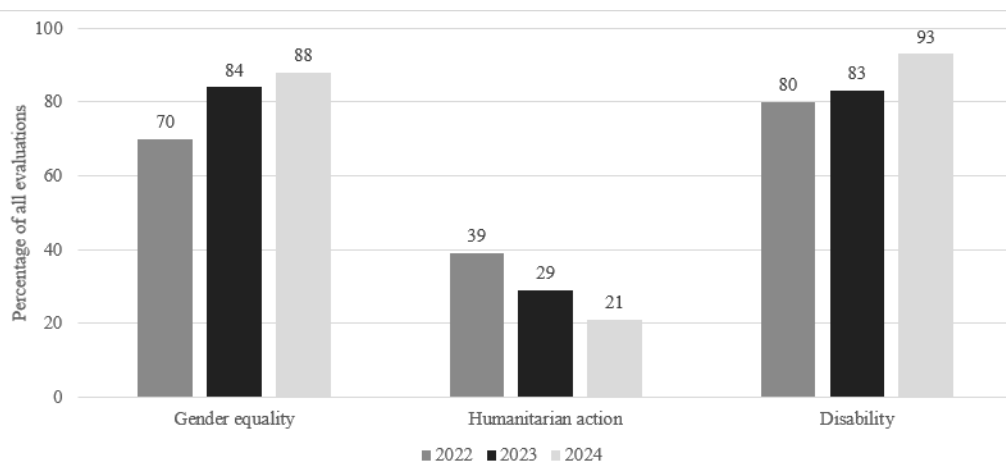
54. UNICEF also supported two country-led evaluations in 2024: the evaluation of the Strategy for Exercising the Rights of the Child 2019–2023 in Montenegro; and the evaluation of the nutrition programme 2019–2023 in Timor-Leste.

### 3. Evaluation coverage of disability, gender equality and humanitarian action

55. As figure VII conveys, 2024 saw steady improvement in the percentage of evaluations systematically incorporating gender equality and the inclusion of persons with disabilities.

Figure VII

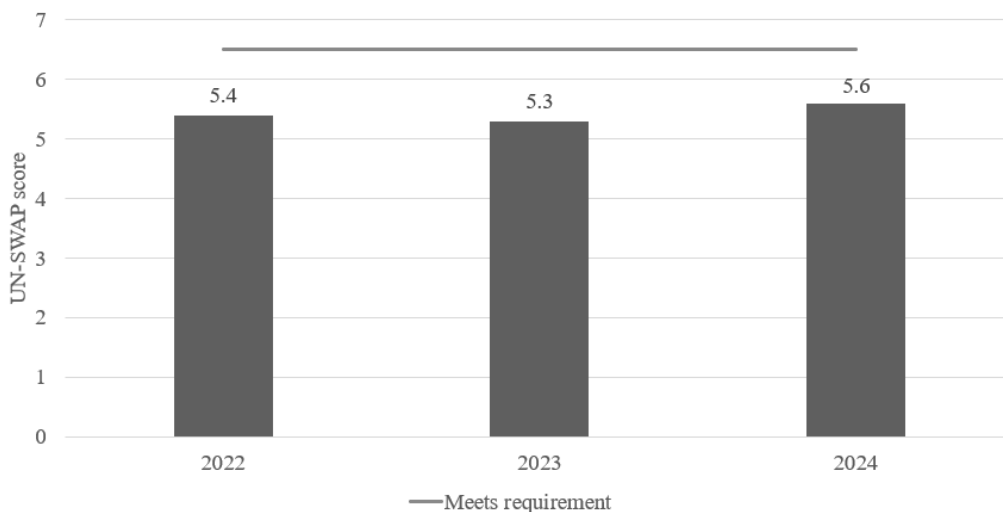
#### Evaluations covering disability, gender equality and humanitarian action, 2022–2024



Source: GEROS.

56. By contrast, progress in meaningfully incorporating gender considerations into UNICEF evaluations was insufficient, with only a modest increase in the System-Wide Action Plan on Gender Equality and Women's Empowerment (UN-SWAP) score to 5.6 in 2024. This lacklustre progress comes after a significant investment of time and effort in 2023 and 2024 to improve performance in this area, e.g. through focused trainings and webinars and the establishment of gender focal points throughout the function. Efforts will continue in 2025 and beyond until substantial and sustained improvement is observed.

Figure VIII  
Evaluation performance under the United Nations System-Wide Action Plan on Gender Equality and Women's Empowerment, 2022–2024

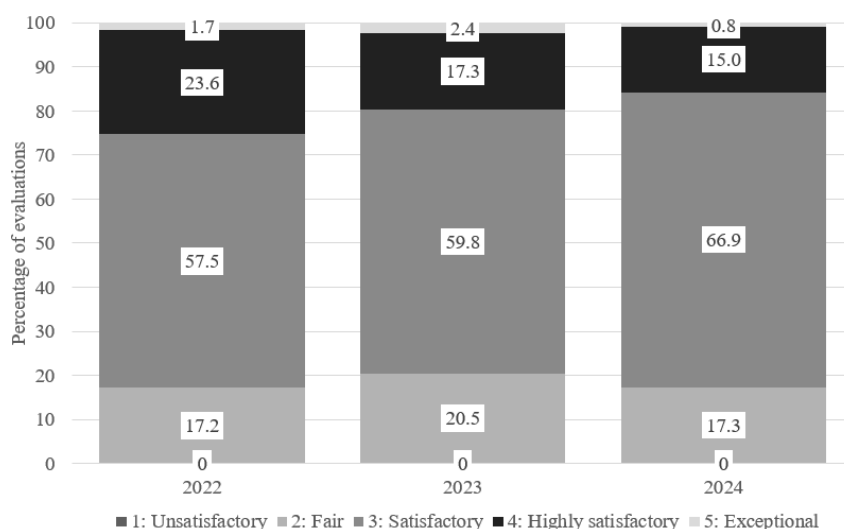


Source: GEROS.

### C. Quality of evaluations

57. Quality remained satisfactory to highly satisfactory overall, improving slightly in 2024. There was a significant increase in evaluations rated “highly satisfactory” or “satisfactory”, the proportion of evaluations rated “fair” decreased and no evaluations were rated “unsatisfactory”. Within these broad patterns, quality remained uneven. In 2024, the overall GEROS ratings ranged from 68.9 to 83 (out of 100) across regions, with a median quality score of 3 (satisfactory) on a scale from 1 to 5. The average scores also varied across regions, ranging from a low of 2.6 to a high of 3.3. These variations highlight the need for tools that ensure consistency in evaluation practices, greater emphasis on quality over quantity, sharper prioritization, better use of GEROS data in talent and performance management processes, and strengthened quality control and quality assurance processes.

Figure IX  
Overall evaluation quality ratings, 2022–2024



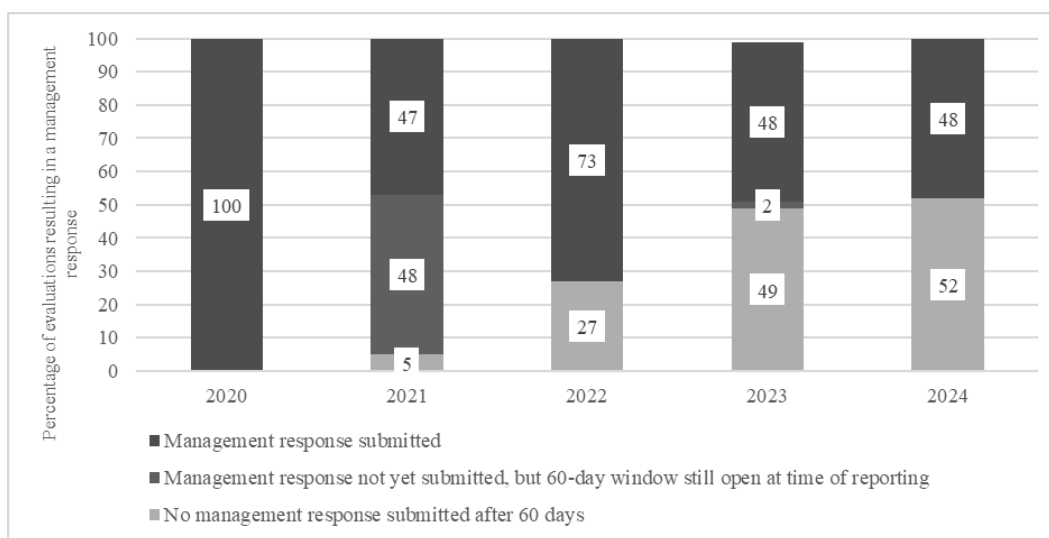
Source: GEROS.

## D. Evaluation management response submission and implementation

58. As shown in figure X, timeliness of management response submissions in the past two years stagnated at about half of the evaluation reports, a significant decline from almost two thirds in 2022. The evaluation function will continue to support management, at all levels of the organization, to submit management responses within the 60-day window, as required by the evaluation policy.

Figure X

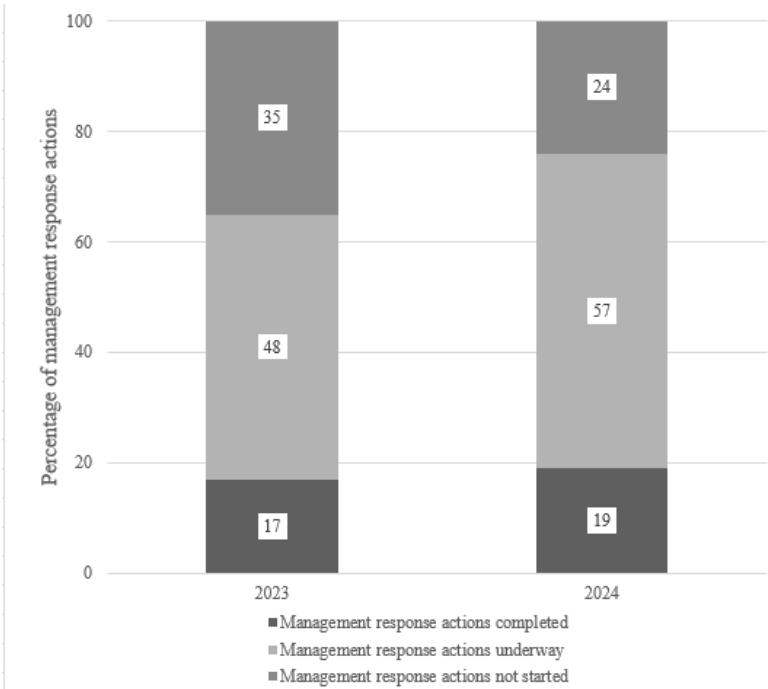
### Timeliness of compliance with management response requirement, 2020–2024



Source: EISI database.

59. By contrast, among those management responses submitted, there was a positive trend in management's initiation of and follow-through on actions. Furthermore, root cause analysis revealed that uninitiated or cancelled actions were largely owed to challenges in the extensive coordination, long-term commitment and external collaboration entailed in engaging meaningfully with government partners and other stakeholders.

Figure XI  
Management response actions and implementation status, 2023–2024

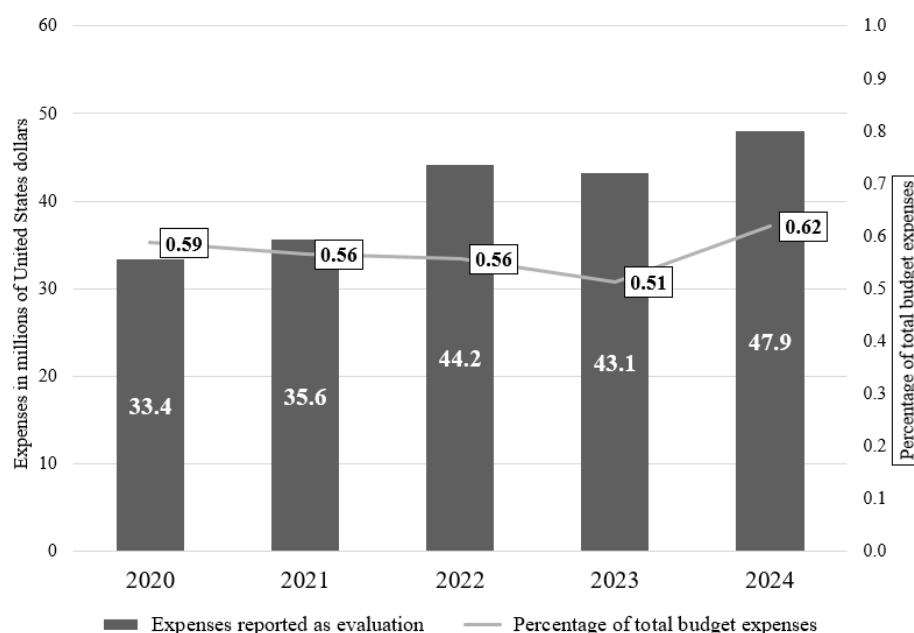


Source: EISI database.

**E. Financial resources and expenditure reported as evaluation**

60. Programme expenditures on evaluation saw a positive increase in 2024, both in absolute terms (thanks to the Executive Board’s approval to raise the Evaluation Pooled Fund ceiling, and the subsequent additional allocation to the Fund by management) and as a percentage of total programme expenses (owing to a decline in total programme expenses, which decreased from \$8.5 billion to \$7.8 billion).

Figure XII  
**UNICEF expenditure reported as evaluation, in millions of United States dollars and as a percentage of total expenses, 2020–2024**

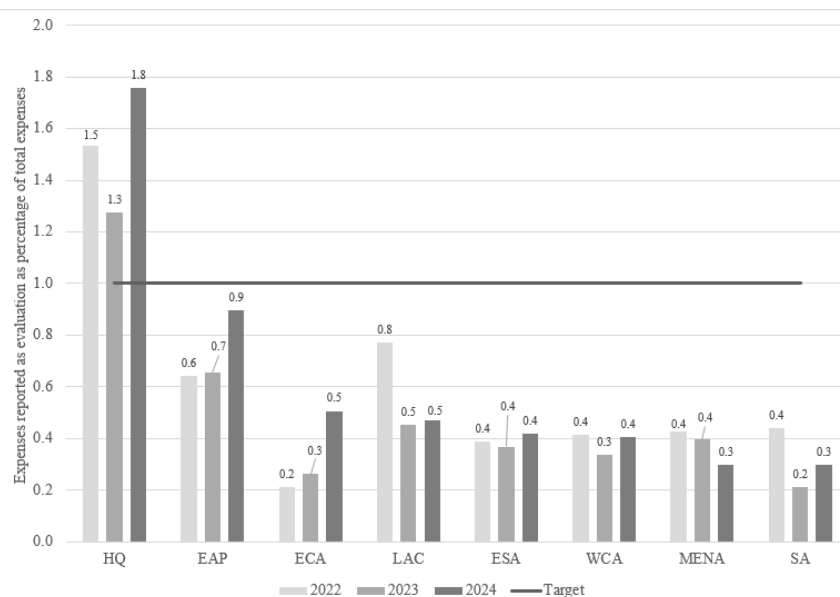


*Note:* The bars and the amounts included inside show the expenses reported as evaluation in millions of United States dollars for each year. The line and numbers denote percentages of the total expenses reported as evaluation expense per year.

*Source:* Data for the numerator: expense reported as evaluation; data for the denominator: total net programme expense; extracted from VISION, 28 March 2025.

61. The Evaluation Office was the only unit that surpassed the 1 per cent organizational target for evaluation expenditure, owing to a drawdown of Evaluation Pooled Fund resources that had been allocated for 2025 in order to meet 2024 work plan commitments. Importantly, a significant proportion of these allocations was invested at the decentralized level, namely by strengthening government capacity (i.e. through ExCEL and IMPRESS) and by funding 16 positions at the regional and country levels. At the decentralized level, almost all regions successfully reversed the trend of declining expenditure witnessed between 2022 and 2023, thanks in large part to the “soft-wiring” elements of policy implementation whereby the evaluation leadership and regional and country leadership worked together over the course of the year to troubleshoot underinvestment. Despite these efforts, however, only one region (East Asia and the Pacific) came close to achieving the 1 per cent target, and expenditure levels remained far below the target elsewhere (i.e. between 0.3 to 0.5 per cent); while two regions (East Asia and the Pacific and Europe and Central Asia) have shown steadily increasing investment over time, investment in the other regions has remained relatively static.

Figure XIII  
Expenditure reported as evaluation as a percentage of total expenses, by region, 2022–2024



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: VISION, extracted on 5 March 2024.

62. Accuracy has improved for all expenses not directly related to evaluation, except for planning and programme reviews, which doubled as a percentage of “inaccurate” coding and in dollar terms. The Evaluation Office put in place ongoing training and monitoring on activity coding, resulting in a significant improvement in accuracy. This said, miscoding still represents 12.2 per cent of total expenses.

Table 2

**Activities reported as evaluation expenditure, by activity type, 2023 and 2024**

Activities	2023 expenditure (in United States dollars)	Percentage	2024 expenditure (in United States dollars)	Percentage
Activities directly related to evaluation	35 163 181	81.5%	42 071 555	87.8%
Data and surveys	1 546 505	3.6%	221 046	0.5%
Planning and programme reviews	580 956	1.3%	1 259 217	2.6%
Monitoring	935 320	2.2%	864 701	1.8%
Research	185 091	0.4%	158 623	0.3%
Evidence generation	372 498	0.9%	364 328	0.8%
Systems strengthening	2 417 720	5.6%	1 829 042	3.8%
Other	1 945 589	4.5%	1 154 821	2.4%
<b>Total</b>	<b>43 146 861</b>	<b>100.0%</b>	<b>47 923 330</b>	<b>100.0%</b>

Source: Activities reported as evaluation expenses retrieved from VISION, 6 March 2025.

## IV. Conclusions and way forward

63. A series of positive shifts were observed in 2024, beginning with the launch of the revised evaluation policy and its numerous measures to optimize the function for maximum positive impact on the work of the organization. In addition, 2024 witnessed a welcome stabilization or reversal of trends in several key performance indicators; this progress was owed to a concerted effort by the evaluation function

and UNICEF management, begun in 2022 and continuing in 2023 and 2024, to work together to address areas requiring improvement.

64. The year also yielded numerous examples of the function's significant and tangible impact on the work for children. On one level, a range of touchstone evaluations at all levels of the organization demonstrated high value – including an evaluation of the implementation of the UNICEF Strategic Plan, whose findings and recommendations management systematically and meaningfully incorporated into the draft of the forthcoming Strategic Plan. On another level, 2024 saw a significant scale-up of strategically crucial initiatives – impact evaluation, NECD and evaluation syntheses – that are collectively working to “move the needle” for children at the systemic level, well beyond the immediate sphere of influence of UNICEF itself.

65. Taken together, the advances of 2024 underscore the unique value proposition of evaluation as an investment – a small investment, targeted at just 1 per cent of the organization's total programme budget – that is helping to future-fit the organization for maximum performance, achieve higher return on investment for children, manage its risks, and enhance partners' trust in the organization. Crucially, none of these gains was the sole result of the evaluation function's efforts, but rather of close collaboration between the function and UNICEF management.

66. In 2025 and beyond, this work to sustain and build on the positive momentum of 2024 will need to continue and expand. Collaboration across all levels of the function, and between the function and management, will be scaled up to ensure the continued strengthening of the function and its enabling environment. Particular attention will be placed on achieving greater consistency of performance across the function, strengthening the focus on outcome-level change, and harnessing strategic communications to enhance evaluation use and build the evaluation culture. Three concrete tools to complement these efforts – i.e. the policy implementation strategy, policy procedure and evaluation handbook – will be completed.

67. A draft Plan for global evaluations, 2026–2029, will also be developed for Executive Board consideration in 2025. Towards this end, a more strongly whole-of-function and more robustly risk-based system will be put in place to enhance the role of independence and impartiality in evaluation planning, aid in the prioritization of the most critical evaluation topics, and maximize the efficiency of evaluation efforts across the three levels of the organization.

68. All of these efforts will take place within a wider context of unprecedented uncertainty and change for UNICEF, beginning with an organization-wide Global Review in the first half of 2025. Adapting to these challenges, the function will additionally focus on:

- (a) assisting the organization as it navigates these changes by providing management with relevant evaluation inputs into the change process;

- (b) safeguarding the integrity of the function, ensuring that it remains robust, that its independence is preserved (and, in keeping with previous Executive Board decisions, strengthened), and that its unique status as an oversight function with a dual learning and accountability role remains intact;

- (c) continuing the effort to ensure that the function is adequately and predictably resourced despite – and arguably in light of – current financial constraints, including through the replenishment of the Evaluation Pooled Fund and the retention of its human capacity at all three levels;

- (d) reorganizing the function itself, identifying ways to create efficiencies and cost-efficiencies within the wider organizational change process.

69. In keeping with Executive Board decision 2023/20, in late 2025 an independent evaluation will be commissioned to assess the implementation of the revised evaluation policy, and specifically its independence-strengthening measures.

## **V. Draft decision**

*The Executive Board*

*Takes note* of the annual report for 2024 on the evaluation function in UNICEF ([E/ICEF/2025/23](#)) and its management response ([E/ICEF/2025/24](#)).

---