



Economic and Social Council

Distr.: General
18 April 2022

Original: English

For decision

United Nations Children's Fund

Executive Board

Annual session 2022

14–17 June 2022

Item 7 of the provisional agenda*

Annual report for 2021 on the evaluation function in UNICEF

Summary

The present report provides an overview of the UNICEF evaluation function in 2021. It outlines the challenging operational context encountered by UNICEF in light of the coronavirus disease 2019 (COVID-19) pandemic, and the adaptations undertaken to continue to meet the learning and accountability needs of the organization and its stakeholders. It presents an overview of the performance of the evaluation function in 2021, as well as an analysis of achievements and lessons learned throughout the 2018–2021 quadrennium.

The report concludes with forward-looking reflections on the evaluation function and the lessons learned, and specifically what these mean for how the evaluation function can ensure that it is fit for purpose in the next quadrennium in support of the United Nations Decade of Action to deliver on the Sustainable Development Goals.

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

* E/ICEF/2022/9.

Note: The present document was processed in its entirety by UNICEF.



I. Overview

1. In 2021, as UNICEF continued to face the dual challenge of the coronavirus disease 2019 (COVID-19) pandemic and the continued erosion of progress in achieving the Sustainable Development Goals, so too did the evaluation function. Despite numerous significant challenges, the evaluation function adapted and remained strong. The constraints imposed by COVID-19 led to a range of creative solutions, including the increased use of real-time evaluations, real-time assessments and remote data-collection modalities. Moreover, increased expenditure on evaluation contributed to notable gains in meeting the commitments outlined in the evaluation policy of UNICEF ([E/ICEF/2018/14](#)). The number of submissions increased, as did geographical coverage, and the function made use of a more diversified set of products, allowing offices to undertake evaluative exercises despite the challenging circumstances.

2. Evaluation quality also remained high in 2021. At the same time, some aspects requiring continued improvement were identified, including in those that are often most crucial for ensuring evaluation use: executive summaries and lessons learned sections. Moreover, despite the expansion of the suite of evaluative solutions offered in response to COVID-19, the share of non-traditional evaluative exercises, such as impact evaluations, evaluability assessments and joint and inter-agency evaluations, remained comparatively small, as did government engagement in evaluations – all of which are areas that will be crucial in light of the urgency of the United Nations Decade of Action to deliver on the Sustainable Development Goals. Lastly, whereas compliance with the requirement of a management response was high, the timeliness of management responses declined in some cases, and it remains unknown whether high compliance resulted in consistently meaningful action on the corresponding evaluation recommendations.

3. Against this backdrop, it is more vital than ever that the evaluation function be fully attuned to the accountability and learning needs of the organization, generating evidence to strengthen the organizational response and help UNICEF and partners to accelerate progress for all children everywhere. The key takeaways of this report therefore offer important lessons that will inform the way forward in the 2022–2025 quadrennium. First, it will be necessary to place ever-greater emphasis on outcome- and impact-level evaluations that help UNICEF and partners measure the end results of UNICEF work on child rights and progress towards the Sustainable Development Goals. At the same time, the evaluation function must produce more rapid, light-touch exercises that can help to maximize the likelihood that UNICEF and its partners will succeed in the first instance. Second, sustained evaluation quality cannot be taken as a given; vigilance and creative solutions are required if quality is to remain high. Third, increased attention must be paid to organizational follow-through on evaluations, including management responses that are timely, robust and actively implemented and monitored. Lastly, adequate resourcing of the function will require continued attention if the organization is to meet the commitments made in response to Executive Board decisions in this regard, if progress in meeting the performance targets set for the function is to continue, and if the strategic direction set for the function is to succeed.

4. These and other measures will be vital in ensuring that the evaluation function is able to evolve, improve and keep pace with unprecedented challenges in support of the Decade of Action. An independent external peer review of the evaluation will be

commissioned in 2022 that will further explore these areas, and, in so doing, inform the revision of the UNICEF evaluation policy in 2023.

II. Introduction

5. The year 2021 began on extremely challenging footing. With most of the Sustainable Development Goals off track and the world still coping with the coronavirus disease 2019 (COVID-19) pandemic, the urgency of the United Nations Decade of Action to deliver on the Sustainable Development Goals has been thrown into sharp relief, with children's rights under threat to an extent not seen in generations. Hard-won progress has been interrupted, and in some cases eroded, as a result of the pandemic. At the same time, protracted armed conflict, food insecurity, rising levels of poverty and climate change-induced natural disasters have continued to pose significant threats to the realization of children's rights and the acceleration of progress against the Goals. Against this backdrop, it is more vital than ever that the work of the evaluation function be fully attuned to the accountability and learning needs of the organization, generating evidence to strengthen the organizational response and help UNICEF and partners to accelerate progress towards realizing the rights of all children everywhere.

6. Despite the many challenges posed by the pandemic in 2021, there were a number of key accomplishments. One was an increase in absolute expenditure in evaluation, with an additional \$15 million spent in 2021 compared with 2018. Increases were especially notable in the Middle East and North Africa region and the South Asia region, which saw increased investment both in terms of absolute expenditures and as a percentage of programme expenditure. Geographical coverage of evaluations also improved in 2021, particularly in the Middle East and North Africa region, where coverage had historically been difficult to attain.

7. There were also indications of improvement in the quality of evaluations. Quality scores in the Global Evaluation Reports Oversight System (GEROS) steadily improved in terms of evaluation methodology and findings, while the quality of recommendations remained strong. These advances enhanced the credibility and utility of evaluations, as shown by increased compliance in the submission of evaluation management responses, which reached nearly 100 per cent in 2021. Importantly, a range of highly influential evaluations – that is, evaluations on whose recommendations key actions were taken that informed the work of the organization in meaningful ways.

8. In the spirit of the United Nations reform agenda and in line with the quadrennial comprehensive policy review of operational activities for development of the United Nations system (General Assembly resolution [75/233](#)), there was an increase in the number of evaluations managed jointly with other United Nations agencies. This trend included efforts to conduct joint evaluations at the decentralized level.

9. These achievements are in many cases the culmination of positive trends that developed over the 2018–2021 quadrennium and represent a strong foundation from which to build as the evaluation function continues to mature and evolve. The function is now ready to strengthen its focus on further improving the quality of evaluations and, together with UNICEF management, the quality and timeliness of management responses and the use of evaluation evidence to inform decision-making.

10. A review of key performance data shows that there is further work to be done in these areas. There is a need to build on progress to date, address remaining challenges, and forge a next-generation evaluation function in the 2022–2025 quadrennium. The

end of the 2018–2021 quadrennium thus presents a critical moment to take stock of the achievements and lessons learned over the previous four years and to reflect on how the Evaluation Office and the evaluation function as a whole can improve the quality, timeliness and use of evaluation evidence in support of the United Nations Decade of Action to deliver on the Sustainable Development Goals.

11. The present report reviews the data on the state of the evaluation function in 2021, both in terms of its performance against established key performance indicators and as benchmarked against the first three years of the quadrennium. It provides a brief contextual analysis of the forces shaping the evaluation function and how the function addressed the challenges and opportunities encountered. The report also provides an overview of progress against the key performance indicators, coupled with an analysis that extracts major takeaways for the 2022–2025 quadrennium. Throughout the document, examples of particularly influential evaluations that informed the work of UNICEF in 2021 are highlighted. These “influential evaluations in 2021” showcase exercises that were rated to be of high quality, with a high level of integration of a gender lens, and which resulted in positive and tangible changes for children.

III. UNICEF operational context: a year of challenges and creative solutions

12. The COVID-19 pandemic continued to necessitate adaption of traditional evaluation approaches in 2021 to meet unprecedented challenges. Several evaluations that had been planned were put on hold in order to shift efforts to COVID-19-related evaluations, with a view to generating lessons on how UNICEF had responded to the pandemic and how the organization could better prepare for future health emergencies. For example, the real-time evaluation of the COVID-19 response in Chad informed interventions to reach the most vulnerable people through food distributions and cash transfers, and the real-time evaluation of the UNICEF response to COVID-19 in Gabon informed the development of the new country programme of cooperation in that country.

13. At the same time, numerous evaluations that were not explicitly focused on the COVID-19 response integrated COVID-related lines of inquiry. For example, the scope of the summative evaluation of the Afghan Women’s Leadership Initiative in support of adolescent girls was updated to include an analysis of how key supplies could be provided to prepare adolescent girls affected by COVID-19 to return to school. The evaluation of UNICEF geographic targeting approaches in Mongolia also expanded its scope to assess the impact of the pandemic on children.

14. In prompting the evaluation function to think and act creatively, the pandemic also resulted in a range of evaluative exercises, including real-time assessments and community rapid assessments in 2021. Drawing on previous experience in humanitarian evaluation, these and other creative solutions that were rolled out in light of pandemic-related restrictions demonstrated that it is possible to produce evaluations without undertaking on-site data collection.

15. While these adaptations have proved essential to meeting the learning and accountability needs of the organization in challenging times, they have important limitations. In some cases, reduced access to key stakeholders and rights holders affected the balanced representation of populations that were the hardest to reach and of children at risk. Lack of field access meant that evaluators had limited interaction with the context in which they operated, and were obliged to rely on less-than-ideal

data sets. The evaluation function made every effort to consider the associated risks and limitations and adapt sampling strategies to mitigate shortcomings and unintended effects and biases. In many cases, however, remote data collection represented a workaround to the challenges at hand. While this approach has been crucial to meeting the evaluation needs of the organization, it cannot be a long-term substitute for more traditional and robust methods of harnessing evidence.

Box I

Influential evaluations in 2021: integrating gender into the UNICEF response to the coronavirus disease 2019 (COVID-19)

The COVID-19 pandemic added another layer of barriers that had the potential to further isolate women and girls from information, services and employment. Understanding the extent to which the UNICEF COVID-19 response has accommodated and considered the unique challenges that women and girls face is critical. Using a participatory, co-designed and co-interpreted approach ensured that findings from the “Real-time evaluation of gender integration and effectiveness in the UNICEF COVID-19 response in South Asia” were available to UNICEF country offices and implementing partners in the midst of the response, allowing them to make mid-course adjustments and adaptations.

IV. Performance overview of the evaluation function: ongoing progress and key lessons

16. Analysis of the data produced in relation to key performance indicators over the past year and the 2018–2021 quadrennium as a whole shows that the evaluation function has remained strong in the face of tremendous challenges. Progress is owed in large part to the significant investments made over the course of the quadrennium. Increased expenditure on evaluation led to notable improvements in meeting the commitments outlined in the revised evaluation policy of UNICEF ([E/ICEF/2018/14](#)).

17. The present section analyses performance in terms of evaluation submission, geographical coverage, quality and use, and outlines investments made to strengthen the evaluation function.

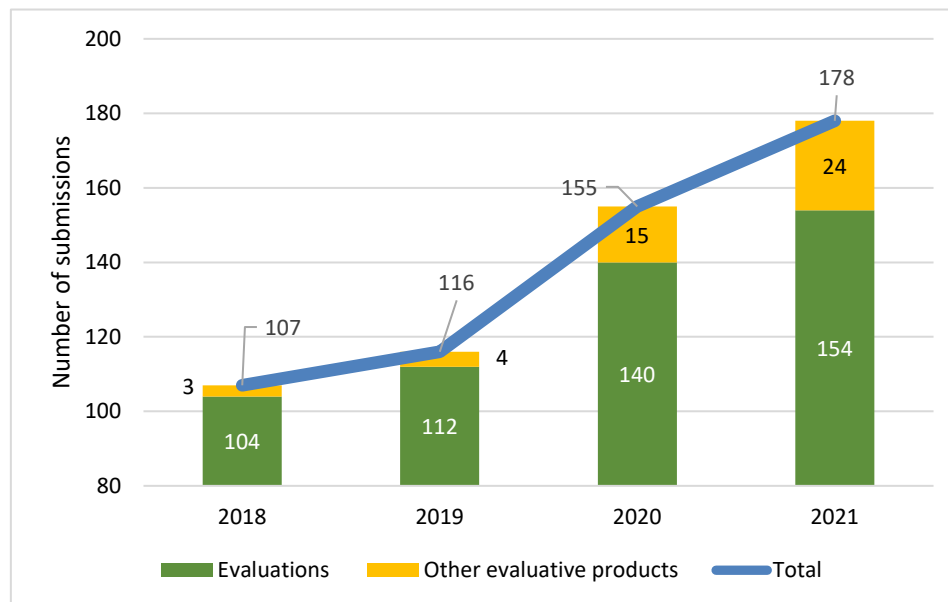
A. Number of evaluation submissions and geographical coverage

18. The coverage norms set out in the revised evaluation policy of UNICEF are designed to ensure that the organization conducts an adequate number of evaluations of programmes, strategies and policies related to the Goal Areas of the UNICEF Strategic Plan, 2022–2025 in order to inform policy, practice and decision-making. In 2021, there was an increase in both the number of evaluation submissions and in geographical coverage. Building on the momentum from previous years, the UNICEF evaluation function produced the highest number of submissions on record.

19. Evaluations in UNICEF are conducted at all three levels of the organization (headquarters and regional and country levels), with the large majority undertaken at country level. Of the 178 evaluative products submitted in 2021, 14 were conducted at headquarters level, while 164 were conducted at the decentralized level. As figure

I shows, 23 more evaluative products were generated in 2021 than in 2020. Throughout the course of the quadrennium, the number of evaluative products increased by 67 per cent.

Figure I
Number of evaluation submissions, 2018–2021



Source: Evidence Information Systems Integration (EISI).

20. Throughout 2021, the COVID-19 pandemic, along with the growing number of Level 3 emergencies, continued to underscore the importance of generating evidence that is not only robust enough to inform decision-making, but is also produced quickly, efficiently and cost-effectively. The evaluation function responded to this need with a range of creative solutions, including the increased use of real-time evaluations and real-time assessments.

21. Six evaluability assessments were conducted in 2021. These exercises sought to enhance programme planning and design by gauging the extent to which programmes would be prepared to be meaningfully evaluated at a later date. In addition, two reviews were conducted that aimed to provide timely feedback on programme implementation of the education management information and monitoring systems. Findings from these reviews will be crucial in programme scale-up. Fourteen real-time assessments were submitted in 2021, compared with 5 in 2020; all of these were related to the COVID-19 pandemic and contributed to the evidence base to inform country and regional COVID-19 responses.

22. This expanded suite of evaluative products was key to ensuring that managers had a range of tools to help them generate evidence to improve programming. With this increased flexibility, however, comes a need for greater clarity around how to be adaptive without sacrificing the independence, quality and credibility of evaluation. Moving forward, there will need to be a clearer distinction between independent evaluative exercises and other types of exercises. There will also need to be clearer guidance around which types of exercises are eligible for funding from the 1 per cent earmarked for evaluation in the evaluation policy and counted towards the 1 per cent expenditure benchmark set out by the Executive Board in its decision 2018/10.

23. As table I and figure II convey, these two top-line trends – increased output and the move to a more diversified set of evaluative products – occurred at all levels of the function in 2021. Submissions by all regions except headquarters and the West and Central Africa region surpassed those of previous years. The highest number of submissions was from the Eastern and Southern Africa region and the South Asia region, while headquarters had the lowest number.

Table I

Number of evaluations and other evaluation products submitted, by location, 2021

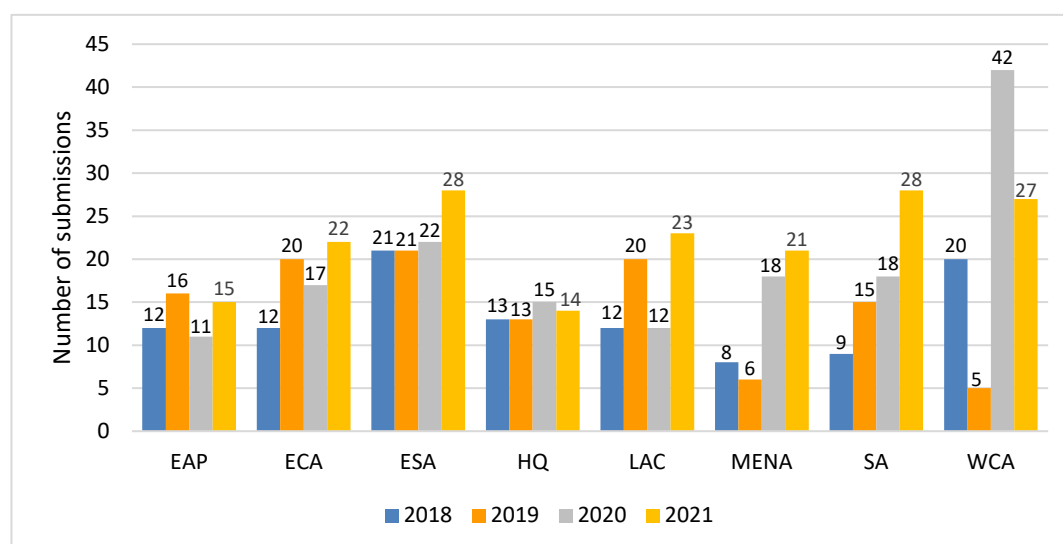
Location	Products submitted				Total
	Evaluations	Evaluability assessments	Reviews	Real-time assessments	
EAP	12	1	0	2	15
ECA	19	1	0	2	22
ESA	23	0	0	5	28
HQ	7	2	2	3	14
LAC	22	0	0	1	23
MENA	20	0	0	1	21
SA	25	1	2	0	28
WCA	26	1	0	0	27
Total	154	6	4	14	178

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.

Figure II

Number of evaluation product submissions by location, 2018–2021

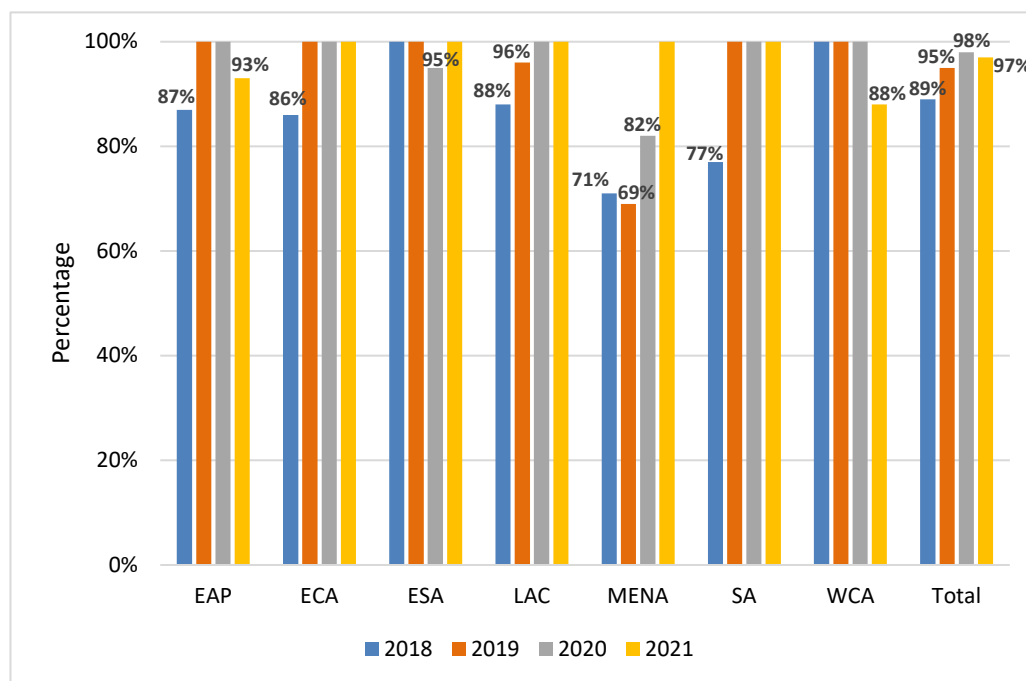


Legend: EAP: East Asia and the Pacific; ECA: Europe and Central Asia; ESA: Eastern and Southern Africa; HQ: headquarters; LAC: Latin America and Caribbean; MENA: Middle East and North Africa; SA: South Asia; WCA: West and Central Africa.

Source: EISI.

24. As shown in figure III, despite the COVID-19 pandemic, there was only a negligible decline in overall geographical coverage (country and regional offices that have conducted an evaluation) in the regions from 2020 to 2021 (from 98 per cent to 97 per cent). This was possible thanks to the use of remote techniques and other workarounds. In addition, multi-country evaluations, whereby offices pool financial and human resources, allowed evaluations to be conducted in contexts where resources were limited. Five of the seven regions attained 100 per cent coverage, and improvements were recorded in the Eastern and Southern Africa region and the Middle East and North Africa region. Coverage declined in the East Asia and the Pacific region and in the West and Central Africa region because one country in the former region and three in the latter have not conducted evaluations since 2019. It is envisaged that the country offices with low coverage rates will conduct evaluations in 2022.

Figure III
Evaluation coverage rates, by region, 2018–2021*



* The lack of a data label above a bar means that 100 per cent geographical coverage was achieved.

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.

B. Evaluating impact and collective contributions towards the Sustainable Development Goals

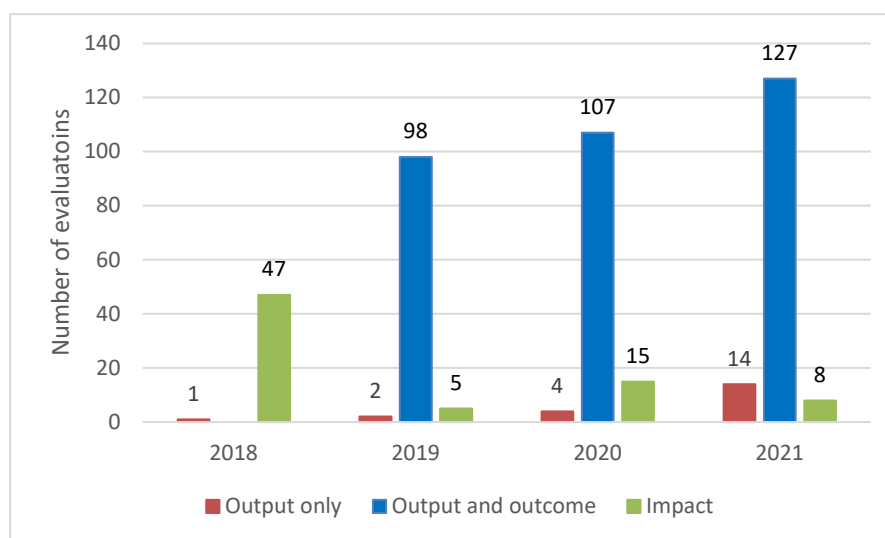
25. While the pandemic highlighted the need for faster evidence generation, it also underscored the need to track progress against the Sustainable Development Goals. A concerning trend in this regard has been the steep drop in the number of evaluations at the impact and outcome levels since 2018, when nearly half of all evaluations included impact-level results. This is due in part to the current use of a more restrictive definition of “impact” compared with previous years, whereby only

experimental designs involving a control group are classified as such. In addition, as the previous GEROS template did not have an option for categorizing evaluations that combine output and outcome, it appears that “outcome” was selected when both levels of results were assessed.

26. Moving forward, impact will be defined according to the level of the results chain at which organizational and programmatic effects are measured rather than according to the specific method used – that is, in accordance with the Development Assistance Committee of the Organisation for Economic Co-operation and Development definition of impact as “the extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects”.¹ Through this emphasis on “higher-level effects” – that is, effects as measured by material, systematic and/or long-lasting change in the lives or experience of those experiencing a given intervention – impact is distinguished from outcomes, which focus on the intermediary or instrumental changes that need to first take place (for example, policy changes enacted and other normative guidance, technical assistance applied, services used) to facilitate the achievement of the ultimate desired impact.

27. In the meantime, of the 154 evaluations assessed, 8 were at impact level, 132 were at outcome and output levels, and only 14 were exclusively at output level. The focus on outputs and outcomes is a trend that has been noticed since 2019. This suggests that there may be a gap in conventional project monitoring, resulting in the need for formative evaluations that look at outputs and outcomes. It is expected that more evaluations will be conducted at the outcome and impact levels, enabling UNICEF to better ascertain the effects of its interventions.

Figure IV
Evaluation focus by level, 2018–2021



Source: GEROS.

28. At the same time, the evaluations produced in 2021 represented a well-distributed mix of backward-looking, summative exercises and forward-looking,

¹ Organisation for Economic Co-operation and Development. “Better criteria for better evaluation: revised evaluation criteria definitions and principles for use” (Paris, Development Assistance Committee Network on Development Evaluation, December 2019).

formative exercises. More than half of evaluations (53 per cent) were summative and formative, while about one third (34 per cent) were formative. Twelve evaluations were purely summative, and there was one meta evaluation. This mix of purposes suggests that both the accountability and learning needs of the organization were being addressed in equal measure.

Table II
Evaluations conducted in 2021, by type

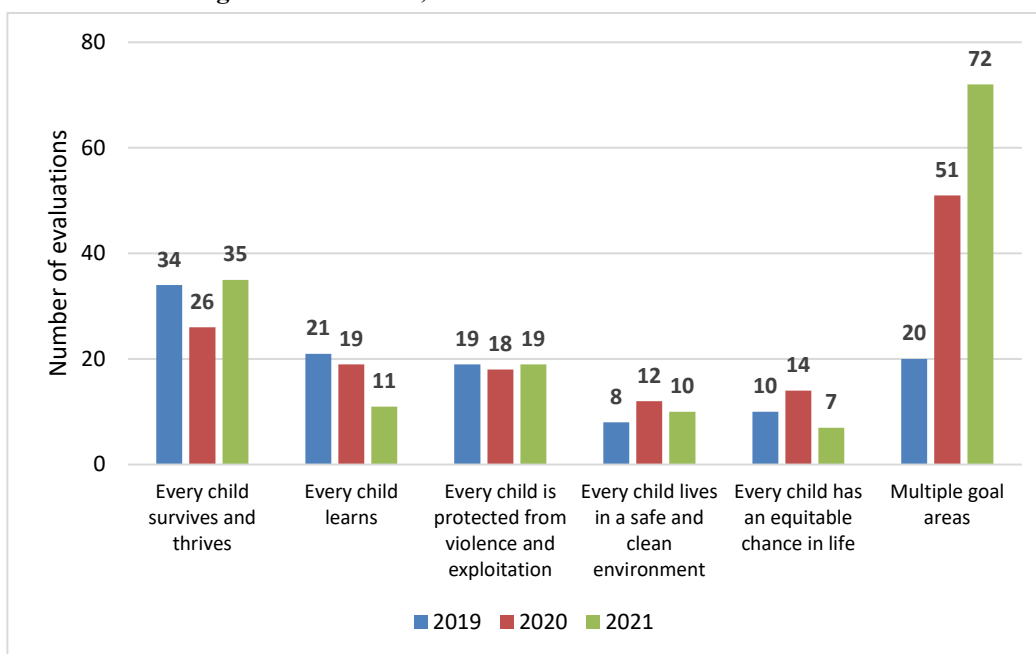
<i>Evaluation type</i>	<i>Number of evaluations</i>	<i>Percentage of all evaluations</i>
Formative	53	34
Summative	19	12
Summative and formative	81	53
Meta evaluation	1	1

Source: GEROS.

1. Goal Area distribution

29. Since 2019, there has been a year-on-year increase in the number of evaluations covering multiple Goal Areas of the UNICEF Strategic Plan, 2018–2021 as regional and country offices shift to more strategic evaluations that inform the entire country programme instead of evaluating single interventions.

Figure V
Thematic coverage of evaluations, 2019–2021



Note: Figures for 2018 are unavailable because systematic tracking of thematic coverage was first introduced in 2019.
Source: GEROS.

30. The plan for global evaluations, 2022–2025 (E/ICEF/2022/3) was designed to ensure broad coverage across all five Goal Areas as well as humanitarian action and cross-cutting themes, and it is therefore envisaged that balance in Goal Area coverage will continue to improve.

2. Joint and inter-agency evaluations

31. Evaluation staff members at all levels of the organization continue to actively participate in the United Nations Evaluation Group and its working groups that seek to strengthen the evaluation practice across the United Nations system. This includes providing input into guidance documents and sharing experiences to foster learning.

32. Beyond this overarching cooperation at a normative level, since 2018, there has also been an increase in evaluations that are managed jointly with one or more United Nations agencies. Of the 10 evaluations managed jointly with other United Nations agencies in 2021, two were at the global level: the inter-agency evaluation synthesis of United Nations system and development bank work towards SDG 6 and the joint evaluation of the UNFPA-UNICEF joint programme on the elimination of female genital mutilation: accelerating change – phase III – 2018–2021.

33. Of particular note is the effort at the decentralized level to conduct joint evaluations in such countries as Côte d'Ivoire, Ghana, Nigeria, the Republic of Moldova, South Sudan, Sri Lanka and Zimbabwe. The evaluation in the Republic of Moldova, for example, was a system-wide evaluation that assessed the collective response of the United Nations system to national development priorities. Positive trends such as these will need to accelerate in the Decade of Action to support efforts to regain lost ground and bring progress on track to meet the Sustainable Development Goals.

34. To date, the practice of issuing joint management responses to joint and inter-agency evaluations has been variable. Evaluations such as these have not always yielded joint management responses, and where these exist, they are often not prescriptive enough, and follow-up has been weak. Positive examples of strong joint management responses do exist, however, particularly in association with inter-agency humanitarian evaluations, which can be mined for best practices. Experience to date has shown that joint management responses are strongest where formal joint/inter-agency accountability mechanisms – such as the Inter-Agency Standing Committee and the Global Action Plan for Healthy Lives and Well-being for All – are in place. This type of action will be important in 2022 and beyond in light of the heightened focus on joint and inter-agency action and in line with the quadrennial comprehensive policy review.

35. Less progress has been achieved with regard to national capacity development for evaluation. Of the evaluations assessed in GEROS, only 4 were led by the host country and 11 were jointly managed with the host country. While this represents a slight increase from previous years, it falls far short of the commitments made in the evaluation policy and will need reinvigorated efforts going forward.

36. The evaluation function continues to support countries in evaluating their progress towards the Sustainable Development Goals. In March and May 2021, for the third year, UNICEF, together with the United Nations Economic Commission for Africa and the Centre for Learning and Evaluation and Results for Anglophone Africa, organized a training for government and UNICEF staff on embedding evaluation in voluntary national reviews in Africa. From 18 to 20 May 2021, UNICEF, along with the Ministry of Foreign Affairs of Finland, the International Institute for Environment and Development, the International Development Evaluation Association, the German Institute for Development Evaluation and EVALSDGs, provided training geared towards government officials from Africa and Latin America on connecting evaluation with national priorities and achieving the Goals.

37. The evaluation function also continued to support a wide range of countries, including Brazil, Cambodia, Fiji, Guinea, India, Kosovo,² Madagascar and Uganda, with evaluations of public policy, evaluation management and the conduct of country-led evaluations.

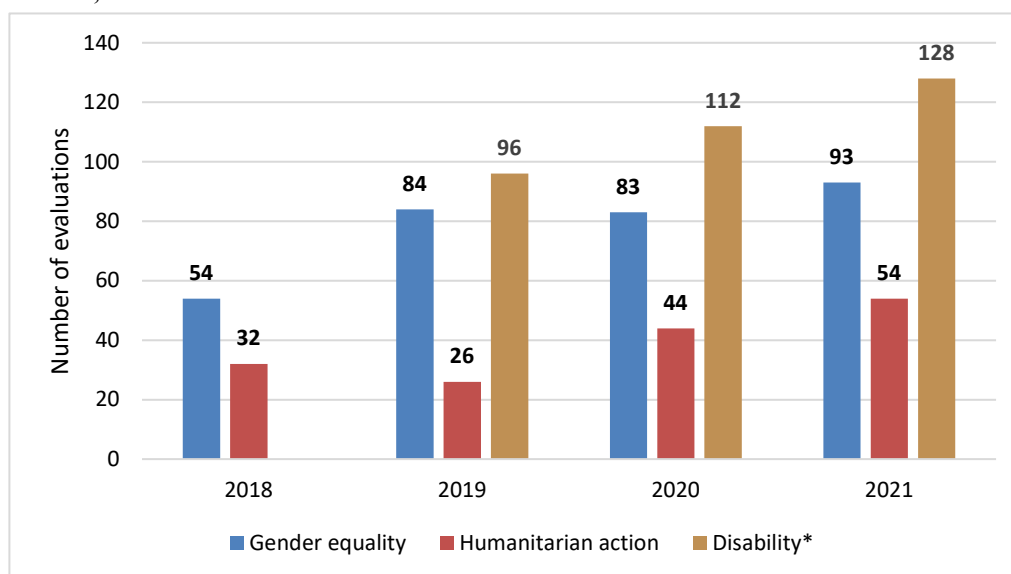
3. Evaluation coverage of disability, gender equality and humanitarian action

38. Evaluation coverage of disability, gender equality and humanitarian action are key focus areas for the function. The number of evaluations covering disability in evaluations continues to increase as offices mainstream disability across all policies and programmes. Evaluation teams are giving specific consideration to the inclusion of disability in evaluation design, findings and recommendations. To ensure that no one is left behind, evaluators conduct focus group discussions with persons with disabilities and/or their families and disaggregate data by disability so that the needs of persons with disabilities are adequately reflected in evaluation reports. The Evaluation Office is undertaking a synthesis of the inclusion of disability in evaluations, which will result in the development of policy guidance for the function. In addition, a cross-sectoral evaluation of UNICEF work on the inclusion of children with disabilities is being scoped.

39. Compared with 2020, there was an increase in the number of evaluations in 2021 covering gender and humanitarian action. Ninety-three evaluations incorporated gender, compared with 83 in 2020. Fifty-four evaluations covered humanitarian action as a cross-cutting theme, compared with 44 in 2020. In humanitarian evaluations, the evaluation function remains committed to being accountable to affected populations. Accountability is ensured through informing affected populations about their rights and entitlements and the expected standards of conduct of evaluation teams, as well as involving them as key stakeholders in the development of recommendations.

Figure VI

Number of evaluations covering disability, gender equality and humanitarian action, 2018–2021



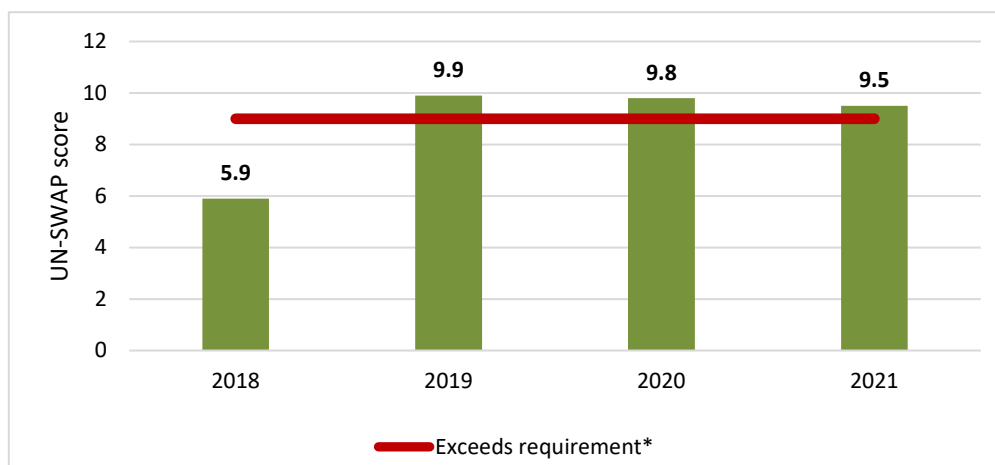
² All references to Kosovo should be understood within the context of Security Council resolution 1244 (1999).

* Tracking of coverage of disability under GEROS started in 2019.
Source: GEROS.

40. Analysis in GEROS shows consistent improvement in the integration of gender equality and the empowerment of women in the scope of analysis of evaluations (evaluation objectives, criteria, questions/indicators in the evaluation framework). Less progress has been made on the use of gender-responsive methodology, methods and tools, and data analysis, and in the reflection of gender analysis in the evaluation findings, conclusions and recommendations. The organization's performance under the United Nations System-wide Action Plan (UN-SWAP) on Gender Equality and the Empowerment of Women remains high. After factoring in the 2019 evaluation of the UNICEF Gender Action Plan, 2018–2021, the overall performance of the 2021 evaluation portfolio "exceeds requirements" even though the overall assessment of the individual evaluations was lower, as shown in figure VII.

Figure VII

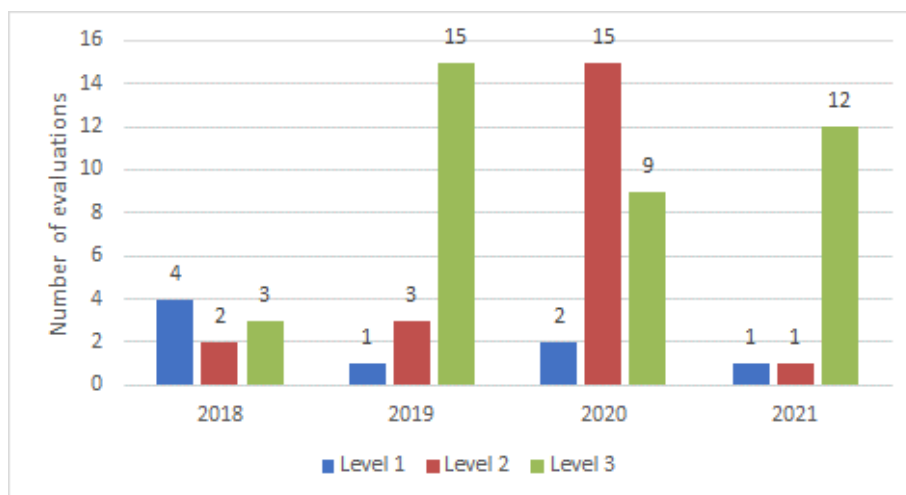
Evaluation performance under the United Nations System-wide Action Plan on Gender Equality and the Empowerment of Women, 2018–2021



* For an entity to be rated "exceeds requirement", evaluation reports must have a UN-SWAP score of 9.01 or above.
Source: GEROS.

41. The function remained focused on the evaluation of declared emergencies in 2021. As figure VIII indicates, 14 evaluations of humanitarian action were conducted in 2021. This included eight evaluations in countries that are currently designated as Level 3 emergencies (i.e., four in Afghanistan, one in South Sudan and three in Yemen) and four evaluations focused on the COVID-19 pandemic. There was also an evaluation of the Level 1 and Level 2 emergencies in Somalia and Cabo Delgado, Mozambique, respectively. Beyond these evaluations, 12 real-time assessments of the COVID-19 response (not included in figure VIII, as they were not predefined as evaluations) were conducted.

Figure VIII
Number of humanitarian evaluations conducted by emergency level, 2018–2021

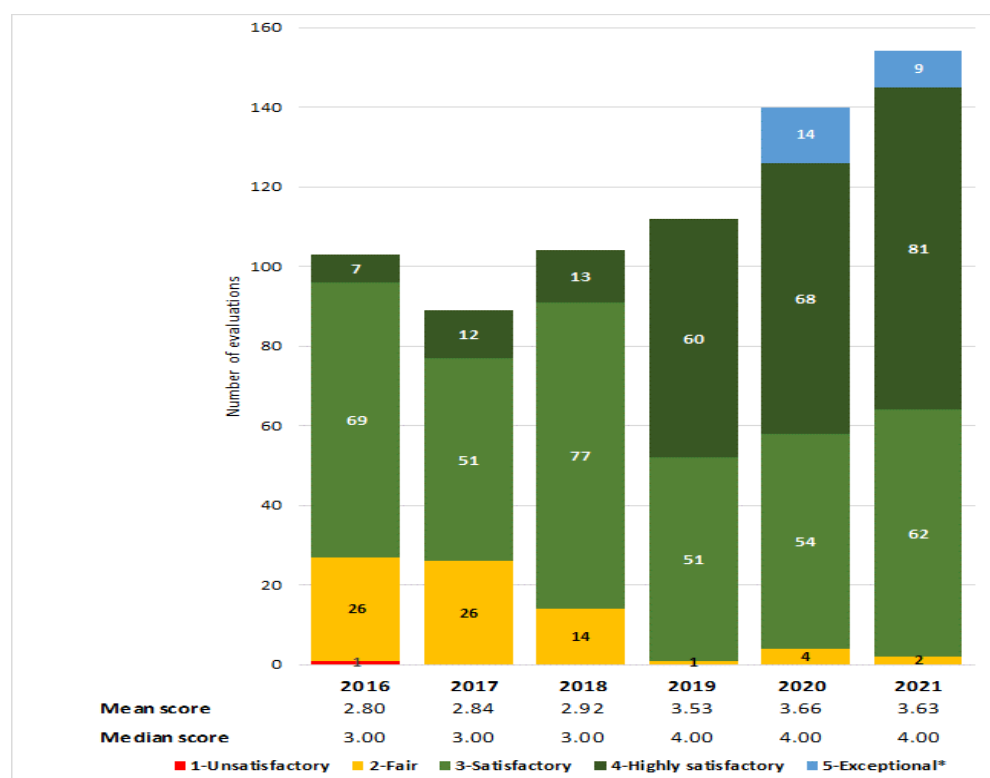


Source: EISI.

C. Evaluation quality

42. Evaluation quality remained high overall in 2021 and, while there was a slight decline in 2020, performance returned to pre-pandemic levels in 2021. Of the 154 evaluations that were independently assessed by an external firm using GEROS, 9 (6 per cent) were rated “exceptional”, 81 (53 per cent) were rated “highly satisfactory” and 62 (40 per cent) were “satisfactory”. Only two evaluations (1 per cent) were rated “fair”. No evaluations received a grade of “unsatisfactory”, a positive trend that has continued since 2016.

Figure IX
Overall evaluation quality ratings, 2016–2021



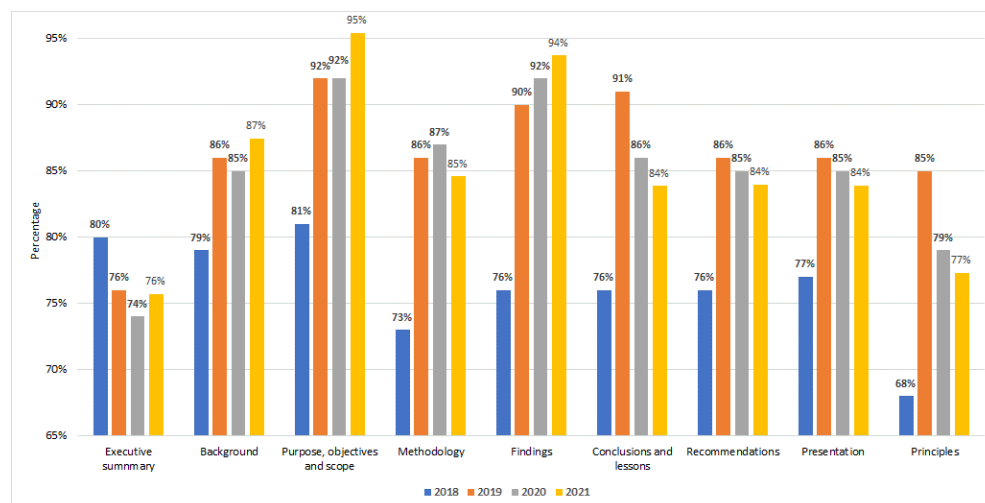
* The “exceptional” category was introduced in the 2020 cycle following a GEROS review.

Source: GEROS.

43. A more granular analysis of the individual elements of evaluation reports reveals a mixed picture. The lessons learned, executive summary and evaluation principles sections of some 2021 evaluations did not perform as well as other aspects of the reports. In this regard, it is worth noting that the lessons learned section, which accounts for 5 per cent of the weighting in GEROS, is now considered a mandatory element. This change was introduced in 2021 to underline the importance of well-formulated lessons for organizational learning. In addition, evaluation quality was also affected by executive summaries that are too long or lack sufficient detail to serve as stand-alone documents. In addition, some reports were not explicit in indicating how disability had been incorporated into the evaluation; this factor, which is a recent dimension included in GEROS, resulted in lower scores under evaluation principles.

44. In many cases, the inability to conduct on-site evidence-gathering has compromised the quality and credibility of evaluations. For example, GEROS data for the 2021 portfolio showed that evaluations were not as strong in methodology compared with the previous two years.

Figure X
Average evaluation quality rating score by report section, 2018–2021



Note: The quality score for each report section ranges from 0 to 100 (where 0 corresponds to the lowest-possible average quality rating and 100 corresponds to a perfect average quality rating among all reports included each year).

Source: GEROS.

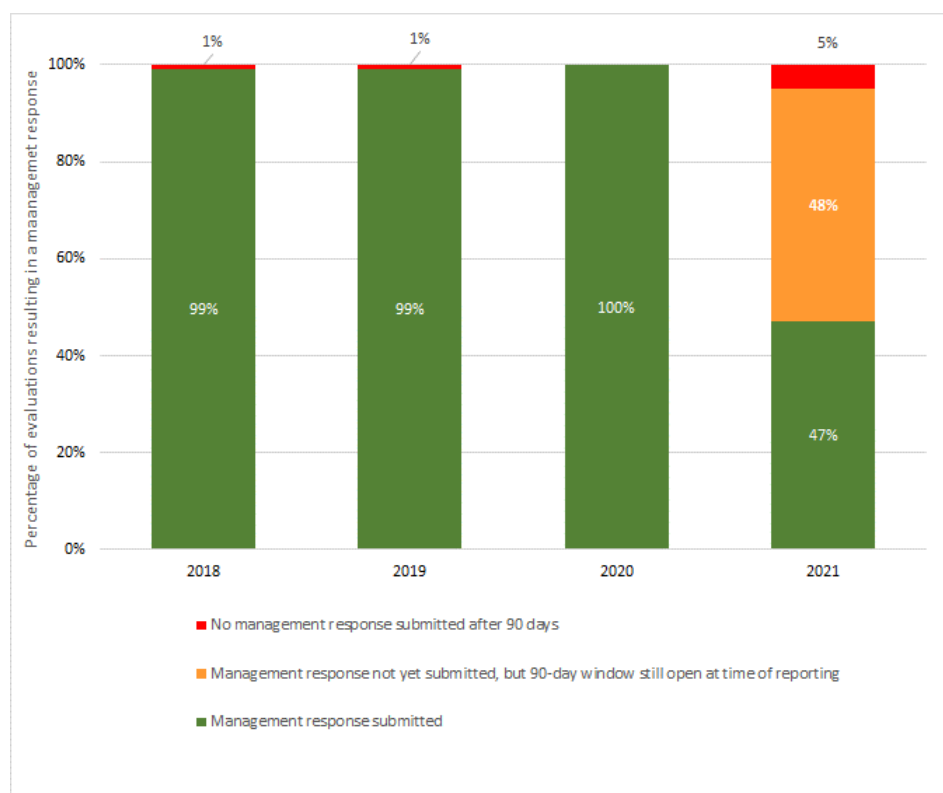
D. Ensuring organizational learning from evaluations

1. Evaluation management response submission and implementation

45. Management responses are an essential platform for ensuring the implementation of evaluation recommendations. The UNICEF evaluation policy requires that management responses be submitted within 60 days following the posting of an evaluation in the Evidence Information Systems Integration (EISI) database. While the time frame was extended in 2020 and 2021 to 90 days due to the COVID-19 pandemic, it has been reinstated to 60 days in 2022.

46. Overall, UNICEF offices are submitting management responses on time. As many as 99 per cent of evaluations conducted between 2018 and 2020 have a management response. For the 2021 cycle, management responses have already been submitted for 73 evaluations (47 per cent). Of the 81 evaluations that do not have a management response, only 7 have passed the 90-day window and are now overdue.

Figure XI
Timeliness of compliance with management response requirement, 2018–2021*

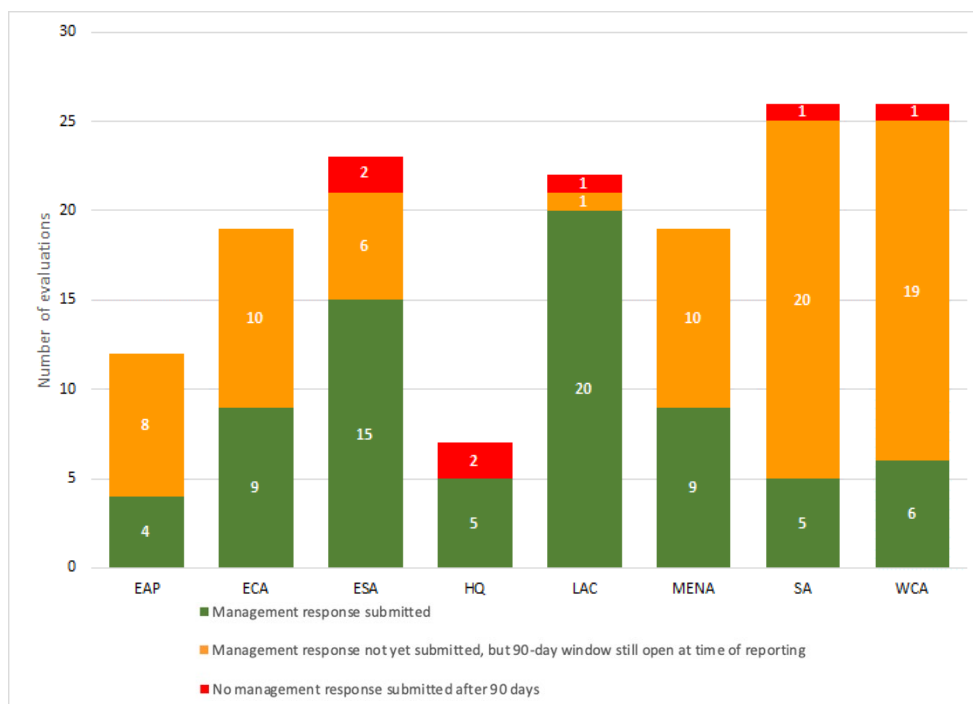


* The time frame for submitting a management response was extended to 90 days from 60 days in 2020 and 2021 because of the COVID-19 pandemic.

Source: EISI.

47. A disproportionate number of overdue management responses occurred at the global level, while three of the seven regions (East Asia and the Pacific, Europe and Central Asia, and the Middle East and North Africa) had no overdue management responses.

Figure XII
Timeliness of compliance with management response requirement, by location, 2021

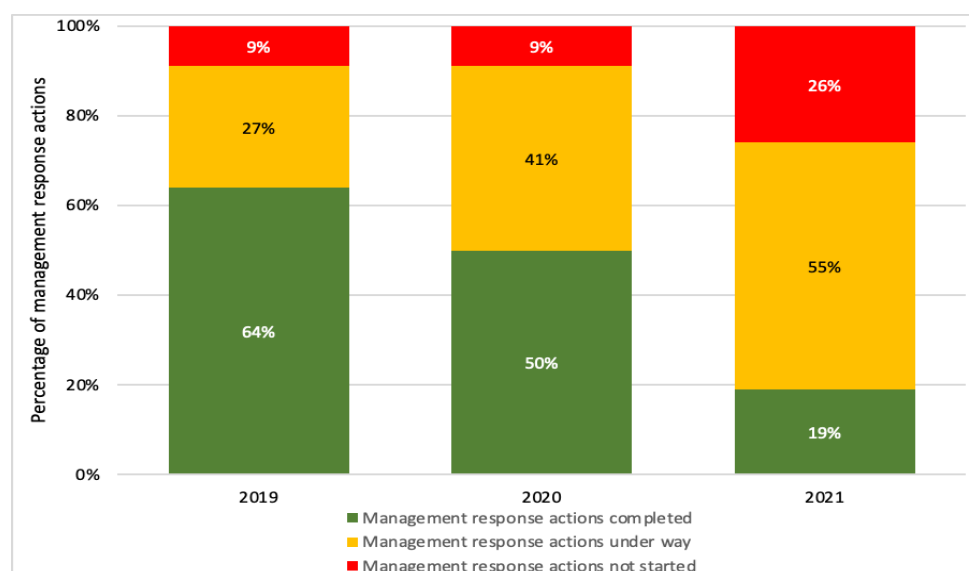


Legend: EAP: East Asia and the Pacific; ECA: Europe and Central Asia; ESA: Eastern and Southern Africa; HQ: headquarters; LAC: Latin America and Caribbean; MENA: Middle East and North Africa; SA: South Asia; WCA: West and Central Africa.

Source: EISI.

48. There has been a significant increase in recent years in the implementation of actions committed to in management responses. Across UNICEF, implementation of management response actions for evaluations completed in 2019 is currently at 91 per cent (64 per cent completed and 27 per cent under way), while only 9 per cent of actions have not yet started. For evaluations done in 2020, implementation is at 91 per cent (50 per cent completed and 41 per cent under way).

Figure XIII
Implementation status of actions identified in management responses, 2019–2021



Source: EISI.

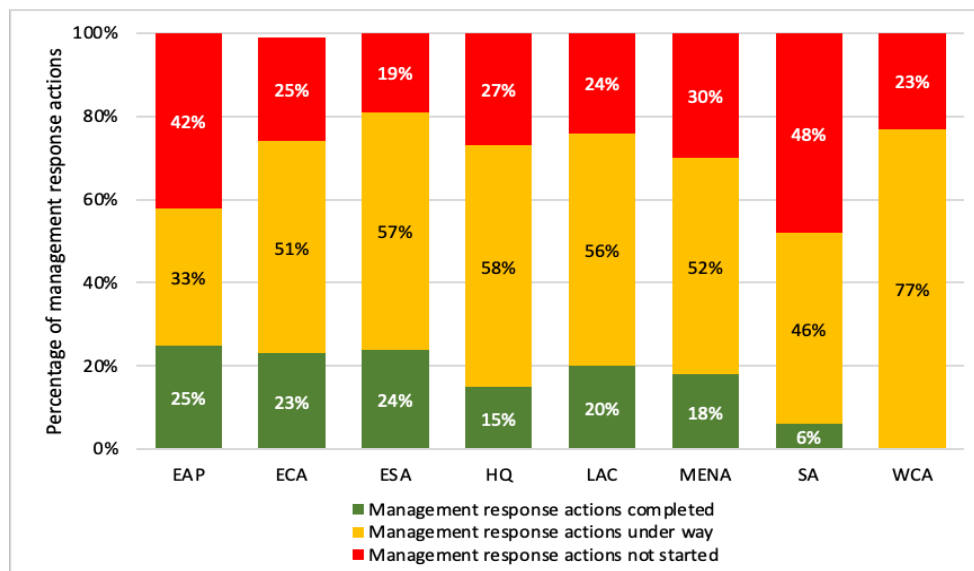
Box II

Influential evaluations in 2021: linking humanitarian and development programming

The “Formative evaluation of UNICEF work to link humanitarian and development programming” informed the UNICEF Strategic Plan, 2022–2025 and the language of the 2020 quadrennial comprehensive policy review, putting an emphasis on “strengthening collaboration, coherence and complementarity” and “contribution to peacebuilding and sustaining peace” internally across UNICEF planning, programming and reporting, and externally through relevant United Nations-wide mechanisms.

49. Regionally, there is marginal variation in the pace of implementation of management response actions for evaluations completed between 2019 and 2021. Implementation was affected by the COVID-19 pandemic, however. It is anticipated that actions will regain momentum in 2022 as pandemic-related restrictions are lifted.

Figure XIV
Implementation status of actions identified in management responses, by location, 2021



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.

Box III

Influential evaluations in 2021: improving programme design

The evaluation of the UNICEF country programme of cooperation in Peru, 2017–2021 informed the design and implementation of the 2022–2026 country programme of cooperation, helping to formulate programmatic strategies and define implementation priorities at the national and subnational levels. The programme design clearly articulates the integration of emergency management into the country programme using lessons learned from the COVID-19 pandemic. In addition, the evaluation was used to account for the use of resources to the Government of Peru, Peruvian society, national partners, cooperation bodies and international donors.

2. Knowledge management and dissemination of evaluation results

50. Even the highest-quality evaluations are of little use if they are not readily accessible to decision makers and easily translatable into practical learning for a wide range of stakeholders. Moreover, high compliance with management response processes does not necessarily mean that management responses are meaningful or well targeted to address the corresponding recommendations. Moving forward, there is a need for fresh thinking within the function, as well as in management, to ensure that evaluations are achieving maximum positive impact for children in the Decade of Action.

51. Ensuring the quality and implementation of management responses is the accountability of UNICEF management. However, there is much that the evaluation function can do to ensure that management response compliance is not only high but

also meaningful. As noted in the previous section, the quality of the two areas most crucial for ensuring evaluation uptake – namely, the executive summary and lessons learned sections – has deteriorated since 2019. The Evaluation Office will seek to strengthen its templates and guidance for these areas as well as for management responses themselves. Beginning in 2022, the Evaluation Office will also commission a review of current management responses and their implementation in order to provide management with a critical and useful appraisal. In the meantime, UNICEF has enhanced the EISI platform to include notifications for outstanding management responses and actions approaching the deadline. In addition, before any management response is closed, documentary evidence must be attached to prove that the actions have been implemented, with a reviewer outside the concerned office verifying that these actions have been fully implemented.

52. It is important to recognize that other parts of the organization, including staff, partners, the Executive Board and the Audit Advisory Committee, also serve as crucial actors in following the management response process and further strengthening overall organizational accountability. In the current climate of information overload, it is therefore crucial that evaluations be communicated as effectively as possible. Towards this end, the Evaluation Office has expanded its partnerships and technological enhancements to allow for easier access to evaluation evidence. The Global Development Commons, launched in September 2020, is a publicly accessible platform where evaluative evidence is accessed through a public domain. The UNICEF-supported platform provides evaluative evidence as its core content, alongside accompanying external partner evidence on children and young people. The Evaluation Office is represented in the UNICEF digital platforms working group, the object of which is to provide digital solutions for data, analytics and knowledge management.

53. The evaluation function also maintains an evaluation resource hub, where lessons learned, technical documents and other tools related to producing an evaluation can be found. Similarly, the evaluation function maintains organization-wide collaboration sites, which are housed in SharePoint. These collaboration sites serve as co-creation spaces as well as digital reference libraries for documents, articles and all materials related to evaluation webinars and learning events.

54. In 2021, the Evaluation Office produced a series of learning events that included webinars (for internal and external audiences), global learning events (two events in 2021), a multimedia COVID-19 learning week, and various Global Development Commons events that brought together UNICEF, universities, non-governmental organizations and students from the Global South. The Global Development Commons webinar series included experts from within UNICEF, as well as external partners including the International Labour Organization, the United Nations Sustainable Development Solutions Network, the Sustainable Development Goals Youth Academy and universities from India, Kenya, Pakistan, Peru and the United States of America.

Box IV

Influential evaluations in 2021: strengthening child protection policies through evaluation

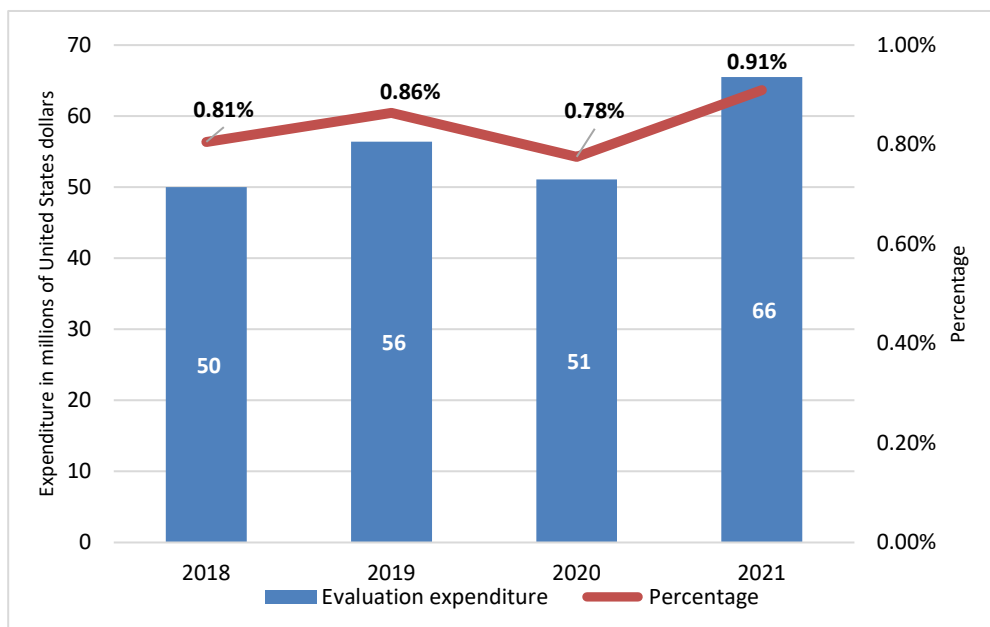
The Children’s Act, 1998, is pivotal to the protection of children’s rights in Ghana, spelling out the role that social workers play in handling maintenance cases. This, however, resulted in social workers being overloaded with cases. The “Formative and summative evaluation of Government of Ghana/UNICEF Child Protection Programme (2012–2019)” outlined that the family tribunal system should be used instead. As a result, the intersectoral standard operating procedure for child protection and family welfare was amended and child maintenance cases are now considered “low risk”, meaning that they should not be managed by social workers. Additional amendments to the Children’s Act have also been proposed because of the evaluation.

E. Financial resources and evaluation expenditure

55. From 2018 to 2021, UNICEF recorded an increase in overall programme spending (from \$6.2 billion in 2018 to \$7.2 billion in 2021). As shown in figure XV, this translated into an increase in overall expenditure on evaluation. Figure XVI provides an overview of trends using the current formula, which was used during the 2018–2021 quadrennium.

Figure XV

UNICEF expenditure on evaluation, in dollar terms and as a percentage of the total programme budget, 2018–2021



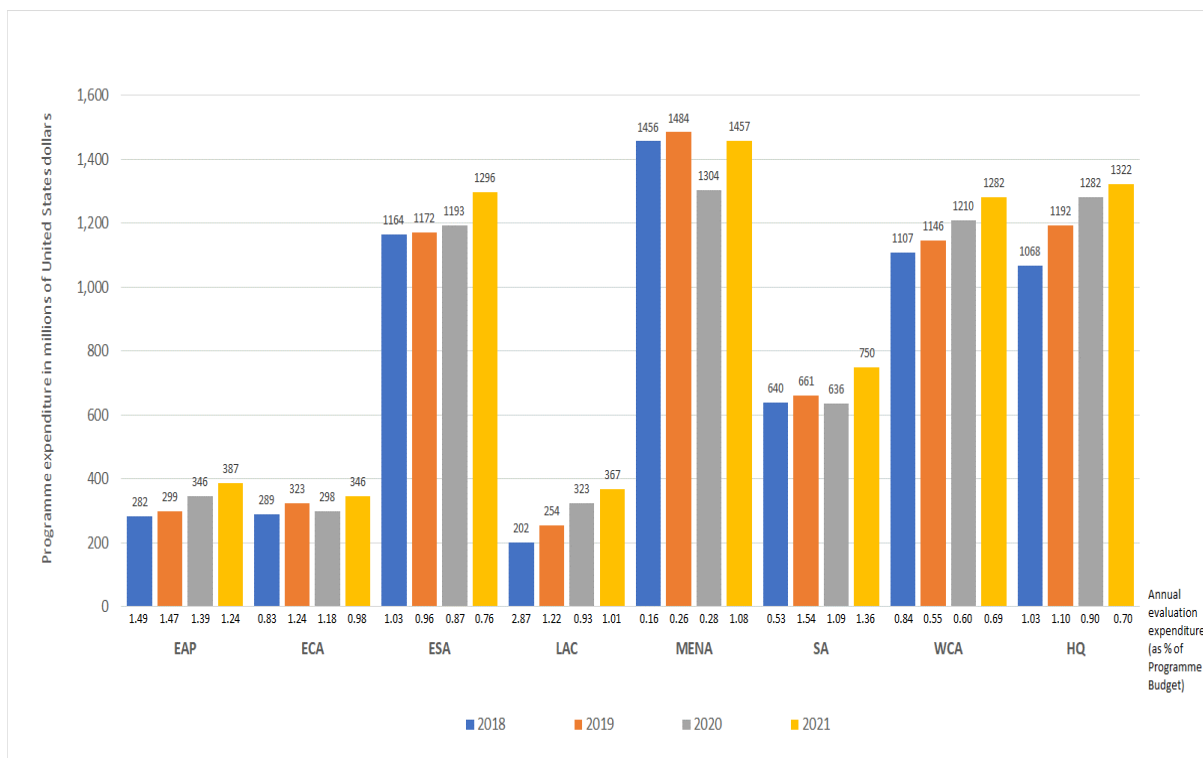
Source: Insight Evaluation Function Performance Dashboard.

56. Performance on the dimension of evaluation expenditure improved in 2021 compared with 2020, with expenditure at 0.91 per cent of overall programme expenditure. At the decentralized level, all regions surpassed the 0.50 per cent mark. A particularly significant increase in spending was recorded in the Middle East and

North Africa region, which invested 0.80 per cent of its programme spend for evaluations, while the South Asia region also saw a surge in evaluation spending.

Figure XVI

Programme expenditure and evaluation expenditure percentage, by location, 2018–2021*



*The bars and the amount above the bars show the programme expenditure for each region and year. The number embedded inside each bar denotes the percentage of the programme budget that was spent on evaluation in each region in the year indicated.

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: Insight Evaluation Function Performance Dashboard.

57. In addition to the need for continued attention to ensuring that the evaluation function is adequately resourced, in the current quadrennium it will be important to benchmark the progress of tracking evaluation expenditure at UNICEF with trends and current practices in other United Nations agencies. The Evaluation Office has begun compiling this information to inform this discussion.

58. Much of the progress outlined in this report is in large part owed to the overall increase in absolute spending on evaluation in recent years, specifically at the decentralized level, which was made possible through the establishment of the Evaluation Pooled Fund. Moving forward, it will be vital to ensure the continued sufficiency of this funding source if the gains in coverage made at the decentralized level are to be sustainable.

1. Human resource capacity

59. The financial resources provided to the evaluation function have been instrumental in ensuring adequate human resources to conduct evaluations. All regions have successfully recruited multi-country evaluation specialists who have

been crucial in the delivery of evaluations, particularly in resource-limited settings. Offices served by these specialists can now pool resources to conduct evaluations. To implement coverage benchmarks of the policy, some offices with budgets greater than \$100 million are also ensuring that there is a dedicated evaluation specialist in place, since they are expected to conduct evaluations every year.

Box V

Influential evaluations: multi-country evaluation of the UNICEF early childhood development response to COVID-19 in the Europe and Central Asia region

The “Multi-country evaluation of the UNICEF early childhood development response to COVID-19 in Europe and Central Asia region: Croatia country case study” provided quick, evidence-based insights to make incremental improvements in early childhood development interventions adapted to the COVID-19 response. The evaluation contributed to enhancing the policy environment in Croatia, with a focus on supporting the most vulnerable families and children in all emergency plans and securing government commitment for scale-up. The evaluation also served as a valuable input to the development of the new UNICEF country programme document in Croatia.

60. In keeping with its commitment to continue building staff capacity in evaluation, the Evaluation Office made significant progress in 2021 on the blended evaluation learning programme for staff and partners, which includes facilitated online learning at the intermediate level through an interactive platform, followed by a two-week in-person advanced evaluation training offered by National University of Singapore. A pilot course and two cohorts of the online facilitated course were held in 2021, with participants drawn from four regions (East Asia and the Pacific, Eastern and Southern Africa, Europe and Central Asia, and South Asia). Certificates of completion were awarded to 53 participants who finalized the course. Recruitment is ongoing for the faculty that will offer the programme in French and Spanish in 2022. The National University of Singapore developed a curriculum for the in-person training, which will be delivered in 2022 to about 100 staff and partners. In addition, two online, self-paced courses are undergoing testing before being launched during the second half of the year.

V. Conclusion and way forward

61. The year 2021 witnessed a continuation of the challenges posed by the COVID-19 pandemic. The evaluation function adapted to these unprecedented challenges by finding new ways of working, new evaluative products to help inform the ongoing response, and new evaluation partnerships to tackle the specific areas of evaluation associated with this unprecedented health emergency.

62. Despite the constraints, the evaluation function continued to perform well on established performance indicators: evaluation activity and coverage increased, as did evaluation expenditure. Evaluation quality generally remained high, and compliance with the management response requirement was nearly universal, suggesting that some of the necessary (if not entirely sufficient) preconditions for meaningful action on evaluation recommendations appear to be in place. As the evaluations showcased throughout this report illustrate, evaluations have benefited the work of the organization in important ways.

63. The pandemic presented not only challenges but also opportunities. Remote data collection became a necessary workaround to overcome access restrictions, and real-time evaluations, real-time assessments, syntheses and reviews demonstrated that useful analysis can be produced rapidly and with a relatively light touch. Even if some of these workarounds are not entirely sustainable, they serve as an illustration that evaluation can be streamlined without significantly compromising the independent evaluative nature of the function's work.

64. At the same time, the function's adaption to ongoing challenges poses valuable lessons for the current quadrennium. The plan for global evaluations, 2022–2025 summarizes these lessons and offers a preliminary glimpse of some of the main currents shaping the strategic direction of the function in 2021. This includes a greatly heightened emphasis on outcome- and impact-level evidence generation. It also includes an expansion of the suite of evaluative exercises available to managers to help inform the work of the organization in a timelier and more rapid manner.

65. The plan for global evaluations also charts a course towards greatly strengthened harmonization and integration, both within the function and with partners. This includes increased coherence across evaluation portfolios for a more holistic, whole-of-child approach; more consistent and meaningful integration of gender, disability and other sources of vulnerability to ensure that no child is left behind; stronger coherence across the three levels of the evaluation function; greater coordination and collaboration with distinct but complementary functions, such as monitoring, research and audit; and greater collaboration with both traditional and non-traditional evaluation partners in the form of enhanced joint and inter-agency evaluations related to the Sustainable Development Goals.

66. The performance data presented in this annual report offer an additional source of knowledge from which to draw lessons that shape the way forward. Some aspects of this analysis reinforce the lessons and strategic direction conveyed in the plan for global evaluations, 2022–2025, while other aspects provide further inputs into the strategic direction for the function with new data and information that were not available when the quadrennial plan was developed.

67. Among the most salient lessons is that there remains an imbalance in the use of traditional evaluations versus less traditional types of evaluative exercises such as evaluability assessments, syntheses, real-time assessments and real-time evaluations. Although the evaluation function expanded its repertoire of evaluative exercise to meet the unique demands and challenges of the COVID-19 pandemic, these non-traditional exercises still represent a small minority of the exercises generated in 2021. If the evaluation function is to remain maximally fit for purpose in the Decade of Action, it will be necessary to rebalance its suite of offerings between traditional evaluation and less traditional modalities. On the one hand, it will be necessary to place increased emphasis on outcome- and impact-level evaluations, which the Executive Board affirmed in its decision 2022/5. On the other hand, the function must also include more of the early, rapid and light-touch exercises that are available, such as evaluability assessments, real-time evaluations and real-time assessments, reviews and evidence syntheses. The former modalities will be essential for measuring tangible results for children from an accountability standpoint, while the latter will be essential for helping the organization achieve these tangible results. In all of these efforts, greater focus on joint and inter-agency evaluative work, with a firmer grounding in the Sustainable Development Goals and the 2030 Agenda for Sustainable Development, will be in order.

68. A second lesson is that sustained evaluation quality cannot be taken as a given; rather, vigilance and creative solutions are required if quality is to remain high and, indeed, improve. While quality remained relatively high overall in 2021, in some key areas it declined or levelled off compared with previous years. Whereas the adaptations associated with COVID-19 might explain some of these shortfalls, they do not necessarily account for all of them. In 2022 more granular analysis will be undertaken to identify the precise reasons for variations in quality. Moreover, a review of the GEROS platform will be undertaken to ensure that it is up to date and consistent in its application, as well as optimized to capture the different quality-assurance approaches that might be required for the various types of evaluative exercises now being produced.

69. Quality alone does not guarantee that an evaluation will achieve its ultimate goal of influencing the actions, decisions and overall strategic direction of UNICEF and its partners. Although it is encouraging that compliance with the management response requirement remained high in 2021, this does not imply that management responses are equally robust or that they are being actively implemented and consistently monitored. A third lesson to emerge from this analysis, therefore, is that increased attention must be paid to organizational follow-through on evaluations. In 2022, a critical analysis will be undertaken to identify any “supply-side” areas that might require improvement – that is, aspects of the evaluations themselves, such as the formulation of recommendations, the GEROS approach to rating recommendation quality, evaluation management, the management response process, or other factors. In parallel, this critical analysis will explore potential areas of improvement on the end-user side, such as the platforms and governance structures surrounding management response development and implementation. Finally, a synthesis will be commissioned to identify those areas of the organization’s operations that have been repeatedly targeted by evaluation recommendations but where meaningful action has not yet been taken. In these ways, the evaluation function will join the rest of the organization in the broader shift towards a stronger outcome- and impact-level orientation as set forth in the Strategic Plan, 2022–2025.

70. Undertaking the major strategic shifts described above while sustaining the gains achieved to date will entail resources. In this vein, the present analysis reveals that adequate resourcing of the function – a directive set by the Executive Board in its decision 2018/10 and underscored in its decision 2022/5 – is an area that will require continued attention in the current quadrennium. While the resourcing of the function as a whole must be secured, of particular concern is the sufficient and predictable resourcing of the decentralized evaluation function – the level at which most evaluations are taking place, where the most progress was made in 2021, and where impact against the Sustainable Development Goals will need to be measured. In recent years, the non-staff costs of the decentralized evaluation function (and some of its staff costs) have been covered by the Evaluation Pooled Fund. The integrated budget, 2022–2025 regularizes the Fund at a proposed level of \$21.5 million. Moving forward, it will be necessary to monitor the Fund on an ongoing basis to ensure that this amount is adequate for sustaining the significant gains to date – and for responding to the ambitious work ahead during the current quadrennium.

71. For the evaluation function to continue to evolve and remain fit for purpose, it will be necessary to reach beyond internal modes of analysis such as this annual report or the plan for global evaluations, 2022–2025. To this end, the findings of this report will be further analysed in 2022 with a view to uncovering the root causes of continued challenges that need to be addressed – and root causes of successes that

need to be scaled up more widely throughout the evaluation function. In addition, in 2022 the Evaluation Office will commission an independent external peer review that will further explore the issues described above, as well as the overall maturity of the evaluation function throughout the organization, from an independent, impartial and objective standpoint. This independent peer review will provide further guidance to the evaluation function and, importantly, serve as a crucial input into the revision of the UNICEF evaluation policy, which will be presented to the Executive Board in 2023.

72. Through the numerous measures described above, it is envisioned that the evaluation function will continue to evolve, improve and keep pace with the unprecedented challenges of our times. By turning the evaluative lens on its own work in 2022 and beyond to ensure that it is as timely, relevant, efficient, effective and impactful as possible, the evaluation function will aim to support organizational learning and accountability in ever better ways, equipping UNICEF, its partners and the Executive Board with the evaluative evidence they need to realize the rights of every child, everywhere, in the Decade of Action.

VI. Draft decision

The Executive Board

1. *Takes note* of the annual report for 2021 on the evaluation function in UNICEF ([E/ICEF/2022/17](#)) and its management response ([E/ICEF/2022/18](#));
 2. *Also takes note* of the evaluation of the UNICEF role as cluster lead/co-lead agency, its summary ([E/ICEF/2022/19](#)) and its management response ([E/ICEF/2022/20](#)).
-