

**DAC/UNEG PEER REVIEW OF  
THE EVALUATION FUNCTION OF  
UNICEF**

**FINAL PEER REVIEW DOCUMENT**  
*July 2017*

## Table of Contents

<b>Acronyms and abbreviations</b> .....	<b>ii</b>
<b>Executive Summary</b> .....	<b>iii</b>
<b>1 Introduction</b> .....	<b>1</b>
1.1 Background to the Peer Review .....	1
1.2 Purpose and scope of the Peer Review .....	2
1.3 The Review Panel .....	2
<b>2 Methodology</b> .....	<b>3</b>
2.1 The framework .....	3
2.2 Methodology and process .....	3
2.3 Limitations.....	4
2.4 Structure of the report .....	5
<b>3 UNICEF, a brief overview</b> .....	<b>6</b>
3.1 Mandate and governance.....	6
3.2 Structure .....	7
3.3 Resources.....	7
<b>4 The Evaluation function in UNICEF: organizational and functional set-up and resources</b> .....	<b>8</b>
4.1 The Evaluation Office .....	8
4.2 The Decentralized Evaluation Function .....	10
4.3 Evaluations by headquarters divisions .....	12
4.4 Human resources.....	13
4.5 Financial resources .....	14
<b>5 Main areas of analysis</b> .....	<b>17</b>
5.1 Evaluation Policy.....	17
5.2 Governance arrangements.....	22
5.3 Evaluation planning and coverage .....	25
5.4 Evaluation management.....	29
5.5 Evaluation quality .....	32
5.6 Evaluation follow-up and use.....	37
5.7 External influence, partnerships and positioning of UNICEF evaluation function .....	41
<b>6 Assessment of the Evaluation Function in UNICEF against the three core criteria: independence, credibility and utility</b> .....	<b>45</b>
6.1 Independence .....	45
6.2 Credibility .....	47
6.3 Utility .....	48
<b>7 Conclusions and recommendations</b> .....	<b>50</b>

### Annexes

Annex 1	Terms of Reference for the Peer Review of UNICEF Evaluation Function
Annex 2	Profiles of Peer Review Panel members and consultants
Annex 3	Normative Framework for the Peer Review of UNICEF Evaluation Function
Annex 4	Bibliography and reference documents
Annex 5	List of interviewed stakeholders
Annex 6	Comparison of UNICEF Revised Evaluation Policy against 2016 UNEG Norms and Standards

# ACRONYMS AND ABBREVIATIONS

DEF	Decentralized Evaluation Function
EA	Evaluability Assessment
EB	Executive Board
ED	Executive Director
EO	UNICEF Evaluation Office
EQA	Evaluation Quality Assessment
EvalNet	OECD-DAC Network on Development Evaluation
G staff	General Service staff (G1 the most junior, G6 the most senior)
HQ	UNICEF headquarters in New York
JIU	Joint Inspection Unit
KPI/s	Key Performance Indicator/s
M&E	Monitoring and Evaluation
MOPAN	Multilateral Organisation Performance Assessment Network
MSs	Member State/s
NECD	National Evaluation Capacity Development
PR	Peer Review
P staff	Professional staff level (P1 the most junior, P5 the most senior)
SDGs	Sustainable Development Goals
SP	UNICEF Strategic Plan 2014-2017
ToR	Terms of Reference
UNEG	UN Evaluation Group
UNEG N&S	UNEG Norms and Standards
UNDP	United Nations Development Programme
UNGA	United Nations General Assembly

# EXECUTIVE SUMMARY

## Introduction

- ES 1. The OECD-DAC Network on Development Evaluation (EvalNet) and the UN Evaluation Group (UNEG) maintain a Joint Task Force to support professional Peer Reviews of the evaluation function of UN organizations. Each Peer Review (PR) is intended to identify good practice and opportunities to further strengthen the evaluation function in the agency under review, with a view to contributing ultimately to improved performance in international development cooperation and humanitarian assistance. Sixteen Peer Reviews have been conducted to date.
- ES 2. A first DAC-UNEG Peer Review of the evaluation function at the United Nations Children's Fund (UNICEF) was conducted in 2006, which led to the preparation of the first evaluation policy for the organization, approved by the UNICEF Executive Board (EB) in 2008. In 2013, following another review of the function, the EB approved a Revised Evaluation Policy (REP), which required UNICEF to undertake a peer review of the implementation of the policy after two years. In 2015, the Executive Board called upon UNICEF to contact UNEG for conducting the second peer review of the evaluation function.

## Purpose, organization and methodology

- ES 3. The assessment aimed at analysing the function against the three core principles of Independence, Credibility and Utility. The scope included both the strategic positioning of evaluation in UNICEF and its functioning at an operational level, at the central and decentralized levels. The core assessment question was: "Are the agency's evaluation policy, function and its products: independent; credible; useful and influential for learning and accountability purposes, as assessed by a Panel of professional evaluation peers against the UNEG Norms and Standards and the evidence base?" The Peer Review was to be forward-looking, providing guidance on how the evaluation function can be further strengthened to meet emerging challenges and opportunities both within the UN system and more broadly.
- ES 4. The primary audiences for the Peer Review are UNICEF Senior Management and the Executive Board, as well as the staff of the Evaluation Office (EO) and more widely across the organisation; the report will be made publicly available through the Web sites of UNICEF Evaluation Office and of the UNEG.
- ES 5. Through the Joint OECD/DAC-UNEG Task Force, a Panel of professional evaluation peers was assembled to conduct the Peer Review, comprising five members. Together, they represented UN agencies, a bilateral donor, a foundation, and the independent professional evaluation community. A consultant supported the Panel in its work.
- ES 6. The UNICEF Peer Review Normative Framework was based on the 2011 UNEG Peer Review Framework, which was revised against the new 2016 UNEG Norms and Standards. Within each of the three pillars, evaluation issues were identified around seven main areas: the Evaluation policy of UNICEF; Governance arrangements; Management of the Evaluation Office, focusing on both operational arrangements and leadership and vision; Evaluation planning; Evaluation quality; Evaluation follow-up and use; and External influence, Partnerships and Positioning.
- ES 7. The Peer Review was conducted in line with the UNEG framework and was characterized by extensive and cordial dialogue between the Peer Review Panel members and the staff of UNICEF and of the Evaluation Office. Tools included extensive analysis of background documents, the review of a small sample of UNICEF evaluation reports, and semi-structured interviews with a wide range of stakeholders within the organization and outside. A mission to UNICEF headquarters in early March 2017 allowed direct interaction with UNICEF senior management and EO staff. The draft report of the Panel was shared with the EO and the

Regional Evaluation/M&E Advisers and Specialists, whose factual corrections, comments and suggestions were integrated, as considered appropriate.

Main findings by area of analysis

- ES 8. **Evaluation policy:** the 2013 Revised Evaluation Policy was found to be comprehensive and broadly aligned with UNEG Norms and Standards on a number of issues. Nevertheless, the REP is informed by a vision of evaluation close to the corporate planning and management functions and largely embedded with other knowledge-generation roles, which are not fully in line with the international standards for evaluation.
- ES 9. **Governance arrangements:** the current set-up foresees the Executive Board exercising oversight of the evaluation function, the Executive Director being fully responsible for the function and its resources and the Evaluation Office with a mandate for corporate-level evaluations and monitoring and reporting responsibilities for the Decentralized Evaluation Function (DEF). This leaves a gap in the internal governance of the function in terms of planning, coordination and harmonization of evaluation activities across the organization, with negative consequences on the efficiency and effectiveness of the function itself.
- ES 10. **Resources for evaluation:** UNICEF is seriously committed and is making progress toward allocating 1% of its financial resources to evaluation. Regarding human resources, the number of staff with 'evaluation' in the title is very high, but the actual amount of time dedicated to evaluation at the level of competence sufficient for management, corresponds in total to approximately 30 full-time staff. For both financial and human resources, there appear to be a significant room for enhancing efficiency and effectiveness.
- ES 11. **Planning and coverage:** many evaluations are planned and completed in UNICEF every year that meet local needs for knowledge, and partly accountability. The decentralized planning system however limits the possibility of synergies across evaluation levels and efficiency in the use of resources. Also, a gap of benchmarks and targets, and of monitoring of all existing criteria, do not allow an analysis of the adequacy of the current evaluative coverage of UNICEF work.
- ES 12. **Evaluation management:** this area of analysis focused almost exclusively on the EO, which largely complies with its role and mandate. Interesting changes are emerging in the management model of corporate evaluations, that may be worth pursuing when funding resources for staff positions will be made independent. It also emerged that some improvements are required on the scope, timeliness and quality of corporate evaluations, and on the internal organization of the office.
- ES 13. **Evaluation quality:** the quality of evaluations in UNICEF has improved over time, partly thanks to the real-time quality assurance support provided to the DEF through external contractual agreements when the Regional Evaluation/M&E advisors do not manage to meet the demand for support; and to the GEROS quality assessment of completed evaluation reports. More guidance is still required on the integration of cross-cutting issues and in the formulation of recommendations.
- ES 14. **Evaluation follow-up and use:** there has been a significant improvement over time in the compliance with requirements for Evaluation Management Responses, and serious attention is given to the implementation of accepted recommendations within the expected timeline. Anecdotal evidence also exists of evaluations used for decision-making, advocacy, management. Gaps were noted in lessons learning and knowledge management across the organization, largely due to its decentralized structure.

ES 15. **External influence, partnerships and positioning:** the EO is well known in the international evaluation community for its commitment to National Evaluation Capacity Development (NECD) and its support to international evaluation networks. On NECD, some anecdotal evidence of results was evident; these nevertheless raise important questions on who, what and how this work should be supported. UNICEF experience in partnership development and on NECD are precious assets that should be brought to bear into the on-going debate within the UN evaluation system, about how best to address the Agenda 2030 challenges on country-led evaluations and NECD.

Main findings by the core criteria for the Peer Review: independence, credibility and utility

ES 16. The Peer Review's assessment of behavioural and organizational independence is short of being satisfactory, both for the EO and for the DEF. This is largely due to the design of the accountability framework and to the 'shared' characteristic of the evaluation function in UNICEF, which lower the boundaries of evaluation as an independent function as part of the corporate oversight mechanism. Furthermore, the admirable and strong corporate culture and focus of 'doing for the benefit of children', which was very apparent to the Peer Review, must not become a justification for limiting formal accountability or for evaluation arrangements subject to potential conflicts of interest.

ES 17. Similarly, the Peer Review's assessment of credibility is short of being satisfactory, both for the EO and for the DEF. This is a direct consequence of the low levels of independence of the function. Other limiting factors include the limited professional evaluation experience of the majority of UNICEF staff managing evaluations, the blurred distinction across roles and responsibilities and the lack of transparency of evaluative coverage of the work of UNICEF at all levels.

ES 18. In the current policy framework, the Peer Review assesses the utility of evaluations in UNICEF to be close to satisfactory, though not quite there yet. There was good evidence of the use of evaluations at the different levels, but improvements appeared necessary in the scope, timeliness and quality of the corporate evaluations. Above all, the Peer Review argues that a better definition is required, of what is the most important role of evaluation within UNICEF, and what needs and gaps it must respond to. This should be part of a process leading to the development, in a consultative manner throughout the organization, of an agreed Theory of Change. This will provide a more robust framework against which the utility of the function can more comprehensively be assessed.

Conclusions and recommendations

ES 19. The purpose of the Peer Review was to support UNICEF in identifying the steps required to further strengthen the function so that it is 'fully-fit' for purpose and well-placed to make the best contribution to the work and strategic positioning of the organization, whilst furthering developments in the field of evaluation. The core question was: "Are the agency's evaluation policy, function and its products: independent; credible; useful and influential for learning and accountability purposes, as assessed by a Panel of professional evaluation peers against the UNEG Norms and Standards and the evidence base?"

ES 20. The Peer Review concluded that there is no short answer to this question, because the evaluation function in UNICEF aims at responding and being 'fit for purpose' for two deeply different views, of what is evaluation.

ES 21. On the one hand, at the corporate level, UNICEF understands evaluation as a shared and decentralized management tool, that mostly contributes to evidence-generation together with research, studies and reviews; and whose contribution to the oversight function is not a primary concern. This is made explicit through the REP and the accountability framework for the

function, which is strongly embedded with planning and management roles, and whereby evaluation duties and tasks, including approval of evaluation plans and reports, are compatible with planning and management roles and responsibilities on the subject being evaluated. Furthermore, in this view, a professional evaluation background is not a necessary requirement for managing evaluation tasks, from the overall leadership of the function to the planning and management of the DEF.

- ES 22. On the other hand, stand the UNEG Norms and Standards for evaluation in the UN system, which are also a reference for the international evaluation community. Pillars here are the independence of the evaluation function, as a fundamental condition for its credibility and a necessary element of utility, in addition to other characteristics. To achieve this, evaluation duties and responsibilities must be clearly separate from any planning and management roles. Also, UNEG is pursuing the professionalization of evaluation, with a complete set of competences for evaluators, as well as for those commissioning and managing evaluations.
- ES 23. The reality of the UNICEF Evaluation function is slightly more nuanced than in the description of the two 'schools of thought', and the Peer Review found evidence at various levels of genuine efforts aimed at bridging the differences between the two. These include initiatives to strengthen the independence of the function; significant efforts to enhance the quality of the evaluation products and the utility of evaluations; real commitment to make available sufficient resources to the function to adequately fulfil its mandate. However, meeting the requirements and expectations of both visions is very challenging and the Peer Review considers that many of the weaknesses it identified in its analysis, stem from the tension between the two.
- ES 24. This tension could lead to an impasse, if UNICEF decided that current evaluations already meet the corporate requirements on evaluation and that only minor adjustments are required, without challenging the fundamental nature of the function. Or, the same tension can become the inspiring challenge that leads UNICEF to transform its evaluation system into an independent, credible and useful function that supports the organization to better implement its mandate and achieve its goals, in full respect of its decentralized nature. This option would be more in line with the strong commitment of the organization to its primary and ultimate clients, the children, and would make of UNICEF an innovator also in the field of evaluation.
- ES 25. Trusting that UNICEF will pick up the challenge, the Peer Review developed six major recommendations, each supported by several possible actions, for consideration by the Executive Board, Senior Management, and the leadership of the evaluation function. These address the need for a new evaluation policy that strengthens the independence of the function, with a direct effect on its credibility; the internal governance of the function, its resources, the quality of the evaluation process and the management of the EO. Each recommendation is supported by a brief synthesis of the evidence in the report that led to its formulation.
- ES 26. The Peer Review is aware that six recommendations, and the suggested actions, may appear to be 'many'. Not all will be regarded as acceptable or feasible, for a variety of reasons. However, because evaluation functions are complex systems with many interacting and inter-related attributes and processes, the Panel is reluctant to rank or prioritize the recommendations, and the related actions, as all appear necessary if the ultimate goal is to make the UNICEF evaluation function fully fit for purpose, and respond to the needs of the organization while being aligned with the international evaluation standards.

#### *An independent and decentralized UNICEF Evaluation Function*

- ES 27. The Peer review identified both strengths and weakness of the UNICEF Evaluation function. At the root of the latter, key issues identified include the conceptual underpinning of a

'shared evaluation function' embedded in the REP accountability framework; the dilution of roles and responsibilities in evaluation with other roles and functions; the constrained role assigned to the EO with regard to the internal governance of the function. In the view of the Peer Review, there is no incompatibility between aligning these fundamentals of the function with the international norms on independence and credibility, while adjusting to the decentralized structure of the organization.

**Recommendation 1:** A new policy for the UNICEF evaluation function should be developed, that takes full account of the decentralized structure of the organization and integrates all the requirements for the independence, credibility and professionalization of the function, in line with the 2016 UNEG Norms and Standards, and with the other recommendations and actions proposed by the Peer Review.

***Specific complementary measures for the implementation of Recommendation 1 include:***

- 1.1 Establish a dual reporting line for Regional Evaluation Advisers: a direct administrative reporting line to the Regional Director, who would also ensure financial resources for the position; and a technical reporting line to the Director of Evaluation in headquarters;
- 1.2 Develop adequate impartiality provisions and safeguards for the behavioural independence of staff with responsibility in planning, commissioning and managing evaluations;
- 1.3 Make publicly available on UNICEF's external web site, all Management Responses to evaluations;
- 1.4 Revamp the Global Evaluation Committee as a platform for substantive discussion between the EO and UNICEF Management on, among others: evaluation topics and planning also on strategic management issues; emerging and compliance issues; sharing of key evaluation findings of corporate relevance.
- 1.5 Development of a Theory of Change for the UNICEF evaluation function;
- 1.6 Development of a Policy implementation strategy, in the form of an Executive Directive, for the future evaluation policy, to guide its operationalization.

***The internal governance of the evaluation function***

ES 28. The Peer Review found evidence that the status-quo of the internal governance of the evaluation function: hampers the opportunities for realising synergies that stronger evaluation coordination at regional and central level may provide; undermines the credibility of the evaluation planning process; leads to duplication of efforts and waste of resources; and restricts improved learning from evaluations across the organization.

ES 29. Recommendation 2 should not be intended as undermining the ultimate responsibility of the Executive Director in the evaluation function, rather as a clear delegation of authority to the Director of Evaluation who would directly report to the ED on all evaluation matters.

**Recommendation 2:** UNICEF should assign to the EO, full responsibility for the internal governance of the evaluation function, and adequate resources for fulfilling this role.

***Specific complementary measures for the implementation of Recommendation 2 include:***

The EO should take the lead, with inputs from the Regional Evaluation Advisers and in consultation with other relevant stakeholders in the organization, on the following actions among others:

- 2.1 Revision of the Key Performance Indicators on evaluation coverage at country level and on the budget allocated to the evaluation function to simultaneously achieve; adequate coverage of UNICEF's work; more accountability for the function itself; and more flexibility to adjust to the very diverse contexts and circumstances of UNICEF country and regional offices;

- 2.2 Ensuring that the Global Evaluation Plan includes strategic corporate-level evaluations, as well as evaluation syntheses or meta-evaluations, that draw on country and/or regional level evaluations identified in consultation with the Regional Evaluation Advisers;
- 2.3 Country level Costed Evaluation Plans should be discussed with both Regional Evaluation Advisers and the EO; these CEPs should include to the extent possible evaluations that will feed into regional, multi-country and global evaluations;
- 2.4 Regional Costed Evaluation Plans should be discussed with the EO and include as appropriate, evaluations that will feed into corporate-level evaluations.
- 2.5 Development of a corporate strategy for National Evaluation Capacity Development, taking into account the evolution of the debate on NECD within UNEG.

### Financial resources for the evaluation function

ES 30. Although UNICEF is increasingly making financial resources for evaluations, evidence available suggests that the application of a 'blanket target' will not maximise efficiency in the allocation and use of available financial resources and that a more rational and efficient option would be elevating the goal of achieving the 1% target to the regional level. This would allow the development of evaluation plans of high strategic relevance to be implemented through regional and multi-country evaluations, as currently done in the CEE/CIS region; while remaining sufficiently close to the national contexts to identify the need for useful and mandatory evaluations at country level. This business model should be supported through some re-distribution of resources across countries within the same region; it would require advocacy and negotiations with donors whose regulations allow this type of mechanism.

ES 31. The Peer Review identified potential risks to the credibility and impartiality of the function, due to the funding modality of a significant share of evaluation staff, in both EO and DEF. Two measures could be envisaged for funding evaluation positions, to ensure that the EO and the DEF maintain access to sufficient human resources to meet the evaluation requirements of the organization, while enhancing the independence and credibility of the function. First, in both EO and Regional Offices as many evaluation positions as possible should be charged to the Regular Resources. Second, for all other positions and for covering the costs of regional and multi-country evaluations a sustainable mechanism should be developed to pool resources from both headquarter-based programme divisions and country offices, to be managed by the EO and the Regional Directors, as appropriate and relevant. In this manner, the contributions from both Regular Resources and Programme Resources would accrue to the corporate spending on evaluation, in a more transparent and visible manner.

**Recommendation 3:** UNICEF should establish new modalities for the funding of the evaluation function at all levels.

### ***Specific complementary measures for the implementation of Recommendation 3 include:***

- 3.1 The allocation of 1% of the financial resources of the organization spent on evaluation should be a target at the regional level, to enable a flexible and more efficient use of resources;
- 3.2 All EO and Regional evaluation staff positions should be funded through Regular Resources, or Programme Resources transferred to the EO and to the Regional Directors, managed under their respective direct responsibility;
- 3.3 A sustainable pool funding mechanism should be developed, to leverage resources from headquarters-based programme divisions and from country-offices, to fund evaluation specialist positions that cannot be funded through Regular Resources, and for conducting evaluations at the regional and/or multi-country level;

3.4 The multi-country evaluation specialist model tested in Cambodia, Malaysia and Myanmar, should be replicated across groups of countries where this may prove appropriate and useful to overcome scarcity of resources for evaluation at country-level.

*Human resources for the evaluation function*

ES 32. The Peer Review considers that mitigating measures should be introduced to safeguard the professionalism and impartiality of the EO, including provisions on the rules governing staff rotation in the evaluation function. Furthermore, the evaluation function would be more effective and efficient by establishing a cohort of full-time evaluation advisers and specialists located at the regional level, or at the national level where the national programme size justifies it, or as multi-country evaluation specialists, who would operate with technical reporting lines to the EO, while closely liaising with the regional and country offices to ensure the relevance and usefulness of evaluations.

ES 33. The Peer Review also identified gaps, in relation to the UNEG Norms and Standards, in the REP and in the practice of the process for the selection and appointment of the Director of the EO. Moreover, with a new role for the Office of internal governance of the function as recommended above, it would appear appropriate to modify accordingly the title of the Director, to strengthen the coherence and unity of the function.

ES 34. All of these issues should be part of a long-term human resource organizational plan for strengthening the function.

**Recommendation 4:** UNICEF should develop a Strategic Human Resource plan for the evaluation function. The Plan should also contain a section that defines the provisions for the selection and appointment of the Director of the Evaluation Function

***The Strategic Human Resource plan for the evaluation function should include the following provisions:***

- 4.1 Establishment of the position of an EO deputy director, at D-1 level;
- 4.2 Ensuring that the Director of Evaluation and Deputy Director together provide credible evaluation experience and competence as required by UNEG Norms and Standards and by the UNEG Competencies Framework for UNEG Heads;
- 4.3 Establishment of the position of Regional Evaluation Adviser at the P-5 level in each and all UNICEF Regional Offices;
- 4.4 Through gradual reallocation of resources over time, establishment of teams of full-time evaluation advisers and specialists in each region, at regional, multi-country and national level where justified, by consolidating the resources currently used for M&E specialists at country level;
- 4.5 Development of specific provisions for the rotation of evaluation staff in UNICEF, that allows staff to pursue a career in evaluation whilst maintaining their behavioural independence.

***The Strategic Human Resource plan for the evaluation function should include the following provisions for the selection and appointment of the Director of the UNICEF Evaluation Function:***

- 4.6 The title of the position should become 'UNICEF Evaluation Director';
- 4.7 The Director of the Evaluation Function should be selected and appointed in agreement between the Executive Director and the Executive Board;
- 4.8 External evaluation expertise should be part of the selection panel for the Director, e.g. at the level of UNEG heads;
- 4.9 The Evaluation Director should report directly to the UNICEF Executive Director, on all matters;

4.10 The ToR for the Director of UNICEF Evaluation should include the systematic presentation of all global evaluation reports to the Executive Board, in addition to the Global Evaluation Plan and the Annual reports on the evaluation function as already the case.

### Quality of evaluations in UNICEF

ES 35. Although a steady improvement of the quality of evaluation reports is evident, some challenges remain. More standardized evaluation processes and more guidance thereon are necessary, together with the need to expand the range of methods and tools applied, to improve the integration of cross-cutting issues into evaluations, and to enhance the quality and usefulness of recommendations. The recommended establishment of an external evaluation advisory committee could contribute to support and enhance the quality of UNICEF evaluation strategies and products, at a time when the EO will go through important changes in staffing, mandate and responsibilities.

**Recommendation 5:** All evaluations planned and commissioned by UNICEF, whether by EO, programme divisions, regional and country offices should aim at achieving the same standards of independence, credibility and utility, and align with the guidance and procedures established by the EO. The EO should be given the responsibility to strengthen the quality of the evaluation process in UNICEF, with an active oversight role on the DEF in this respect.

### ***Specific complementary measures for the implementation of Recommendation 5 include:***

- 5.1 Establishment of an external evaluation advisory committee, similar to the one established by the Independent Evaluation Office in UNDP;
- 5.2 Development of standard guidance for quality assurance of the evaluation process;
- 5.3 Establishment of a real-time Quality Assurance mechanism in support of evaluations at country level in all regional offices, under the responsibility and management of the Regional Evaluation Adviser and the oversight of the EO;
- 5.4 Development and/or adaptation of evaluation guidelines and manuals to fit the UNICEF evaluation process, from inception to completion;
- 5.5 Development and/or adaptation of evaluation guidelines and manuals to improve the integration of human rights and gender equality perspectives in evaluations;
- 5.6 Revision of the criteria that differentiate evaluations from other types of assessments and reviews, and consider a more consistent application of the taxonomy in the titles of evaluation reports;
- 5.7 Revision, in consultation with Senior Management and the Regional Evaluation Advisers of the quality standards for evaluation recommendations, to make them more useful to UNICEF managers while maintaining the possibility for strategic, corporate and cross-cutting issues to be captured;
- 5.8 Revision of the timeframes for the implementation and closure of recommendations that address strategic and corporate-wide issues.

### Management of the EO

ES 36. Areas for improvement were identified in the management of the Evaluation Office, in particular on the internal working arrangements to ensure the timely delivery of evaluation reports and to enable a diversity of perspectives in evaluation. The Peer Review is also aware that more changes will be necessary, also in terms of staff capacity, to meet the challenges raised by the stronger internal evaluation governance role entailed in Recommendation 2, and by Recommendation 5 on quality of evaluations. In this, the Advisory Panel recommended above could play an important role in supporting the Director of Evaluation in transforming the EO into an office better fit for its future purpose.

**Recommendation 6:** The Director of UNICEF Evaluation should revise the internal management processes of the EO.

***Specific complementary measures for the implementation of Recommendation 6 include:***

- 6.1 Improving the efficiency of the corporate-level evaluation processes and the timely delivery of all evaluation products by the EO;
- 6.2 Rotation of the EO evaluation advisers and specialists for the management and conduct of evaluations, so as to enable diversity of perspectives and experience in evaluations of the same thematic areas and programmes;
- 6.3 Inclusion of highly reputed specialists in the subject matter of the evaluation, into the evaluation teams responsible for carrying out evaluations commissioned by the EO and the DEF;
- 6.4 Ensuring that all criteria and Key Performance Indicators identified in the Evaluation Policy are adequately monitored and reported upon.

**The future of the evaluation function**

- ES 37. UNICEF, even within the United Nations, is uniquely placed in terms of (i) the extensive global reach it has in 192 countries through country programmes and National Committees, seven regional offices, a supply operation in Copenhagen and offices in Geneva, Tokyo and Brussels, (ii) its highly decentralized operations, (iii) the presence of evaluation staff in many country offices, (iv) strong research capacity and a research centre in Florence, and (v) sizeable mobilization of resources from the private sector. In the broader development world, UNICEF has played a high profile and critical role in advocating for the rights of children and in producing evidence and data used by global and regional organizations. Looking ahead, we feel that:
- ES 38. UNICEF will need to capitalize much more strategically on the comparative advantages it enjoys. An independent UNICEF evaluation function in particular needs to keep pace with a rapidly changing development and evaluation world which, for its sustainability and progress, requires much greater equity, better assurance of rights, and more respect for human dignity.
- ES 39. The future of development will have to, in addition to ensuring effective service delivery, increasingly rely on generating and managing evidence and knowledge to scale effective interventions and on influencing a broad range of players and investors. The significance of financial contributions from UNICEF are likely to decrease even further as nations become richer, and graduate to becoming middle or high income countries. Scale partners such as governments and impact investors will depend on the knowledge that independent and credible evaluations (and research) can provide - of what works, when and where. This has several implications for UNICEF's evaluation function.

***Reputation and Influence***

- ES 40. Externally, the UNICEF EO continues to be well regarded and appreciated by the international evaluation community for its long-standing commitment to regional and national evaluation capacity development, and for the unique leadership role and contributions it has made to the global development evaluation architecture (e.g. EvalPartners, AfrEA, SDGEval, tools on equity focused evaluation, webinars, MyMandE).
- ES 41. The evaluation community looks to UNICEF to continue to play this important role in an increasingly challenging global context for evidence and evaluation, and to be a much-needed champion for evidence of why equity and inclusion matter. These efforts will have to be more systematically assessed, streamlined and strengthened for greater impact on the lives of children.

### ***Strategic leadership***

- ES 42. The EO staff and its Director appear to have been skilful in navigating UNICEF's complex, decentralized system, building trust and relationships to engage in evaluation.
- ES 43. The EO has a window of opportunity now with its new Director of Evaluation to strengthen its leadership role in UNICEF's evaluation function and strategy by optimizing the evaluation synergies across the organization, evolving and responding to the need for more rapid cycle, strategic evaluations, synthesis reviews and cross organizational learning and accountability in addition to the ongoing evaluations requested of the EO.

### ***Evolving skills and capacities***

- ES 44. With strong leadership, there is also a window of opportunity for UNICEF's evaluation function to diversify its technical capacities to embrace evolving evaluation approaches and to be more influential with a wider range of key traditional and non-traditional actors in a post-2015 SDG and innovation era, from development policy players, to investors and private sector players. Three areas are likely to call for greater attention.
- a) UNICEF's evaluation function will need to equip itself better to deal with new technologies for real time data gathering and analyses through the use of new ICT technologies including the use of mobile, big data, predictive analytics, and machine reading technologies;
  - b) UNICEF's evaluation work will need to encourage innovations in methodologies beyond traditional approaches to embrace complexity and systems approaches, stronger focus on rigorous social return on investment analyses (economic modelling of financial and social returns), and greater attention to efficiency and cost effectiveness;
  - c) UNICEF evaluation staff will need to forge important forward looking innovation partnerships with, for example, the UNICEF Innovation Office and other evaluation Innovation Hubs to develop new, faster and more efficient approaches to monitoring and evaluation in UNICEF, the UN system and beyond.

### ***Learning and influence across complex systems***

- ES 45. Looking ahead (and even now), successful development outcomes rely on the actions of many diverse players and investors. No longer can one agency alone achieve global impact. The ability to learn across organizations, new partnerships, and actors will be crucial, and thus UNICEF's evaluation function will need to equip itself with a broader suite of tools and approaches to learn across systems, new players and to synthesize and communicate learning for influence.
- ES 46. UNICEF is likely to be called upon, given its large country-level and regional physical presence, to play a much more strategic leadership role in managing and learning from such partnerships.
- ES 47. Evaluations will be required to increasingly contribute to improving institutional performance, enhancing effectiveness and efficiency of collective actions, improve joint accountability, and utilize common learnings for delivering results.
- ES 48. The EO will have to play a bigger role in drawing on the quality, independence and coverage of the evaluations it conducts to draw lessons from meta-evaluations and cross-organizational learnings. In this regard, four areas where evaluation may provide insights are likely to warrant special attention:
- a) More can be done to optimize learning and strategic influence and to achieve greater synergies among the key evaluation players across the organization and the field;

- b) More needs to be known about the best ways for including and involving children (particularly adolescents) in evaluation. Though UNICEF and other organizations have promoted children's participation, and this is currently being pursued, it remains a neglected area;
- c) Better understanding is needed of combined impacts of 'development' and 'humanitarian' assistance on children and their families disrupted, upended or nearly destroyed by conflict, human-made disasters and wide-ranging, climate-change events and other emergencies.
- d) Ending violence is critical for sustaining development. Little however is known on prevention strategies that work to end violence particularly against children;

***Better positioning and synergies for greater influence***

ES 49. The potential and opportunity exist for UNICEF to better use the Evaluation Function for strategic corporate evaluations and partnerships that look beyond programme sectors and sections to support effective corporate governance, greater synthesis of knowledge and lessons (with research), risk management (with audit), and strategy and leverage of private sector lessons and investments. One example, among many possible topics, would be to use findings from an evaluation of the private sector's contribution to UNICEF in helping to mobilize greater resources and attention to improving the lives of children.

# 1 INTRODUCTION

## 1.1 *Background to the Peer Review*

1. The OECD-DAC Network on Development Evaluation (EvalNet) and the UN Evaluation Group (UNEG) maintain a Joint Task Force to support professional Peer Reviews of the evaluation function of UN organizations. Each Peer Review (PR) is intended to identify good practice and opportunities to further strengthen the evaluation function in the agency under review, with a view to contributing ultimately to improved performance in international development cooperation and humanitarian assistance.<sup>1</sup> Sixteen Peer Reviews have been conducted to date.
2. A first DAC-UNEG Peer Review of the evaluation function at the United Nations Children's Fund (UNICEF) was conducted in 2006.<sup>2</sup> This led to the preparation of the first evaluation policy for the organization, which was approved by the UNICEF Executive Board (EB) in 2008. In 2013, following another review of the function, the EB approved a Revised Evaluation Policy (REP),<sup>3</sup> which required UNICEF to undertake a Peer Review of the implementation of the evaluation policy after two years. In 2015, the Executive Board called upon UNICEF to contact UNEG to arrange conducting a second Peer Review of the evaluation function.
3. The terms of reference (ToR)<sup>4</sup> for the Peer Review of UNICEF's evaluation function established that the assessment would analyse the function against the three core principles of Independence, Credibility and Utility. The closely related core criteria of impartiality and transparency would also be considered. The focus of the PR was to be on the status and performance of the evaluation function, in particular considering to what extent it is fit for purpose, influential with key stakeholders (internal and external) and matched to UNICEF's evolving approach and organization as proposed in its Strategic Plan 2014-2017. The Peer Review was to be forward-looking, providing guidance on how the evaluation function can be further strengthened to meet emerging challenges and opportunities both within the UN system and more broadly.
4. The primary audiences for the Peer Review are UNICEF Senior Management and the Executive Board, as well as the staff of the Evaluation Office (EO) and more widely across the organisation; the report will be made publicly available through the Web sites of UNICEF Evaluation Office and of the UNEG.
5. This is the final report of the Peer Review. The report should be proactively discussed and disseminated within UNICEF and beyond, including to the members of UNEG and DAC Evaluation Network, in their capacity as secondary audience for all Peer Reviews conducted under their framework, for information on issues of evaluation quality and utility. The Peer Review Panel will also provide feedback on the Peer Review process to the Joint Task Force on Peer Reviews to contribute to the further development of the instrument.

---

<sup>1</sup> See *UNEG Framework for Professional Peer Reviews of the Evaluation Function of UN organizations*, 2011, at <http://www.unevaluation.org/document/detail/945>.

<sup>2</sup> See *Peer Review of the Evaluation Function at the United Nations Children's Fund (UNICEF)*, May 2006, at <http://www.uneval.org/document/detail/98>.

<sup>3</sup> See *UNICEF Revised Evaluation Policy*, 2013, at: [https://www.unicef.org/evaluation/files/unicef-evaluation\\_policy.pdf](https://www.unicef.org/evaluation/files/unicef-evaluation_policy.pdf). <http://www.uneval.org/document/download/2428>

<sup>4</sup> See Annex 1.

## **1.2 Purpose and scope of the Peer Review**

6. The main purpose of the Peer Review was to support UNICEF in identifying steps required to further strengthen its evaluation function so that it is fully fit for purpose and well-placed to make the best contribution to the work and strategic positioning of the organization whilst furthering developments in the field of evaluation. The Peer Review was thus expected to provide recommendations to the Executive Director (ED), the Executive Board and the Evaluation Office for improving the quality of the evaluation function generally, and specifically to inform discussions and decisions about the role, positioning, leadership, vision, resourcing (including capacity) and mandate of the Evaluation Office itself, as well as arrangements for evaluation at the decentralized levels.
7. The scope of the Peer Review included both the strategic positioning of evaluation in UNICEF and its functioning at an operational level, at the central and decentralized levels. The core assessment question was: “Are the agency’s evaluation policy, function and its products: independent; credible; useful and influential for learning and accountability purposes, as assessed by a Panel of professional evaluation peers against the UNEG Norms and Standards and the evidence base?”
8. The timeframe for assessment of the evaluation function of UNICEF by the Peer Review was the period 2013-2016, using the Revised Evaluation Policy (2013) as the baseline reference and with specific focus on arrangements and performance at the beginning of 2017. Whenever necessary, developments since the 2006 Peer Review were also considered.

## **1.3 The Review Panel**

9. Through the Joint OECD/DAC-UNEG Task Force, a Panel of professional evaluation peers was assembled to conduct the Peer Review.<sup>5</sup> Criteria for its composition included:
  - relevant professional experience;
  - independence: to avoid any potential or perceived conflict of interest or partiality, the Panel members did not have any close working relationship to UNICEF evaluation function during the period under analysis, that might influence the Panel’s position and deliberations; one of its members worked in the position of UNICEF Regional Monitoring and Evaluation officer, and left the organization in 2013;
  - institutional affiliations: members were drawn from multilateral and bilateral development agencies and from the foundation sector; and
  - geographical and gender balance.
10. The Panel comprised five members, including two representatives of UN agencies, one of a bilateral donor, one of a foundation, and one independent economist and evaluator from a UNICEF programme country. One consultant had the role of supporting the Panel throughout its work, including preparation of background materials and tools, preliminary analysis, and contribution to drafting the report. The Panel worked through teleconferences and emails, in addition to a one-week mission to UNICEF headquarters in early March 2017.

---

<sup>5</sup> See Annex 2.

## 2 METHODOLOGY

### 2.1 *The framework*

11. The UNICEF Peer Review Normative Framework was based on the 2011 UNEG Peer Review Framework, which was revised for the occasion against the new UNEG Norms and Standards, updated in 2016.<sup>6</sup> The framework is grounded on the three pillars of independence, credibility and utility of the evaluation function. Within each of these, evaluation issues were identified, focusing on 1) the independence and credibility of the function; 2) the quality, use and follow up of evaluations across the organization to promote accountability, learning, and improvement; 3) the leadership, and vision shown by UNICEF's evaluation function, including its capacity to adjust to the changing environment.
12. The ToR for the PR identified seven main areas of analysis, namely:
  - A. The Evaluation policy of UNICEF;
  - B. Governance Arrangements;
  - C. Management of the Evaluation Office, focusing on both operational arrangements and leadership and vision;
  - D. Evaluation planning;
  - E. Evaluation quality;
  - F. Evaluation follow-up and use; and
  - G. External Influence, Partnerships and Positioning.
13. Each area of analysis included a sub-set of issues, identified by both the EO and the Peer Review Panel, which were then individually linked to the respective core principle. The resulting matrix provided the basic analytical tool of the PR.

### 2.2 *Methodology and process*

14. The Peer Review of UNICEF Evaluation Function was conducted in line with the UNEG Peer Review Guidance document<sup>7</sup> and was characterized by extensive and cordial dialogue between the Peer Review Panel members and the staff of the Evaluation Office. This open and transparent approach also informed the interaction with all other key stakeholders in UNICEF.
15. The EO, on behalf of UNICEF, prepared the first draft of the Terms of Reference for the PR. The Panel finalized the ToR and prepared the Normative Framework for the Review, following exchanges with the EO Director. The ToR were circulated to all key stakeholders within UNICEF, ahead of the meetings planned with the Panel members.
16. In December 2016, the EO had launched an organization-wide self-assessment of the status and performance of the evaluation function, against UNEG norms and standards. The framework for the Self-Assessment was broadly consistent with the ToR and the Normative Framework for the Peer Review, and its findings and results were shared with the Panel, and are cross-referenced throughout the report.
17. The PR used the following tools and approaches:

---

<sup>6</sup> See Annex 3.

<sup>7</sup> Peer Review Guidance, draft unpublished document prepared by the UNEG Peer Review Task Force, February 2015.

- Desk review and analysis of background documents, Web site pages and on-line information management systems, on UNICEF and its evaluation function;<sup>8</sup>
  - Review of a small, non-representative sample of fifteen UNICEF evaluation reports, issued between 2013 and 2016, and assessment of the reports' quality using the Global Evaluation Reports Oversight System (GEROS) templates applied until 2015 and from 2016 onward; criteria for selecting the sample were the scope of the evaluation, country and region, and year;
  - Semi-structured interviews, face-to-face or via skype or phone, using check-lists prepared for the different categories of stakeholders; in total, the Panel interviewed 67 informants from among UNICEF Senior Management, headquarters senior staff, Regional Directors, Country Deputy/Representatives, EO staff, Regional Evaluation/M&E Advisers and Specialists at the country level, team leaders of UNICEF evaluations, UNEG Heads and partners. The Panel also held one meeting with the Executive Board, with 15 Member States Representatives;<sup>9</sup>
  - In-depth discussions and exchanges with EO and regional M&E staff;
  - Peer-exchange session between the EO staff and the Peer Review Panel, on a list of jointly selected topics; this consisted of almost two-hours of informal and open discussion on various aspects of the UNICEF evaluation function work, drawing on the experience of the Panel Members within their own or other organizations.
18. Finally, the draft report of the Panel was shared with the EO and the Regional Evaluation/M&E Advisers and Specialists, whose factual corrections, comments and suggestions were integrated, as considered appropriate.
19. The Peer Review process was conducted in the period December 2016-June 2017, and was structured as follows:
- i. Preparation (October-November 2016): mobilization of the Panel;
  - ii. Desk-review (December 2016-February 2017): while the Evaluation Office was undertaking a self-assessment, the Peer Review panel held meetings to discuss and finalize the terms of reference and carried out a desk-review of available documents;
  - iii. Data gathering (March 2017): this phase included a main mission by the Panel to UNICEF headquarters, and follow-up work from respective locations, to carry out interviews with Senior Management, the Executive Board, directors in headquarters, regional and country offices, monitoring and evaluation staff in EO, Regional Evaluation/M&E Advisers and Specialists at the country level and in UNICEF headquarters; the Peer Exchange session was also conducted during the Panel's mission to UNICEF headquarters;
  - iv. Report writing (April-June 2017): this phase included the analysis and triangulation of findings; preparation of the draft report, validation across the Panel, sharing with the EO; preparation of the final report;
  - v. Presentation of the Final Report (September 2017): the Chair of the Peer Review Panel will present the final report at the Executive Board, for the members' consideration and discussion.

### **2.3 Limitations**

20. A professional Peer Review of the evaluation function is not intended to be a fully-fledged evaluation that can comprehensively evaluate practices, processes, and outcomes in depth. Time and resources are the main limitation in this respect.

---

<sup>8</sup> See Annex 4, Bibliography and reference documents

<sup>9</sup> See Annex 5, List of interviewed stakeholders

21. The complexity of the evaluation function that emerged from the initial desk-review led to the decision to defer part of the fact-finding phase, to after the Panel's mission to UNICEF headquarters. In practice, however, time available between the mission and the deadlines for presentation of the draft and final reports to UNICEF, was a constraint for canvassing the views of a larger range of stakeholders. For example, no interviews were held with M&E staff working for Programme divisions in headquarters, and only few with evaluation team leaders who also had experience of country-level managed evaluations. This limited the possibility to better validate some of the findings.
22. Another limitation concerned the country coverage through evaluations. The January 2017 UNICEF High Level Structure indicates 117 country offices, some of which are responsible for more than one country, in particular in the Caribbean and in the Pacific. However, the UNICEF Country Office Portal lists 159 countries where UNICEF has a Country Programme document, or that is part of one. As evaluation reports indicate the country, but not the country office that managed the evaluation, for the sake of simplicity, the analysis was done listing evaluations by country, rather than by office. This may have led to some under-estimation of the actual evaluation coverage, as multi-country, regional and global evaluations may have assessed work in countries that did not have any 'country-focused' evaluation.
23. Furthermore, in consideration of time constraints, the Panel also decided that the analysis of a larger sample of UNICEF evaluation reports was not a priority for the Review, all the more so considering that the GEROS mechanism already provides a thorough independent and systematic analysis of the quality of evaluations produced by the organization.

## **2.4 Structure of the report**

24. This report was largely structured according to the Terms of Reference for the Peer Review. Following the initial sections that present the background, the purpose and scope and the methodology of the Peer Review, Section 4 describes the function, at headquarters and decentralized level, including human and financial resources. Section 5 brings forth and discusses the evidence found during the assessment following the seven areas of analysis listed above. Section 6 analyses the extent to which the UNICEF evaluation function complies with the three core principles of the Peer Review; and Section 7 follows, with the overall conclusions and recommendations. Finally, in Section 8, the Panel offers a few brief opinions regarding the future direction of the evaluation function.
25. In addition to the recommendations in Section 7, the report includes some suggestions, in the form of text underlined, on various aspects of the evaluation function. These address issues that would deserve attention for further strengthening the evaluation function, but do not hold such importance as to be included among the recommendations proposed to UNICEF Executive Board, Senior Management and Evaluation Office.

### 3 UNICEF, A BRIEF OVERVIEW

26. An exhaustive description of UNICEF's mandate and work falls outside the task of the Peer Review. Thus, this section focuses on those features of the organization that have, or may have, relevance for its evaluation function and impact on its scope, efficiency and effectiveness.

#### 3.1 *Mandate and governance*

27. The Mission Statement for the organization, adopted by its Executive Board in 1996, confirms the mandate by the United Nations General Assembly (UNGA) to UNICEF that includes advocating for the protection of children's rights, helping to meet their basic needs and expanding their opportunities to reach their full potential. The Statement also identifies the Convention of the Rights of the Child (CRC) as the guiding framework for the organization, and affirms the corporate commitment to advocate for and promote the survival, protection and development of children, with special attention to the most disadvantaged children, both in peaceful and emergencies context. Over the years, UNICEF's work expanded beyond health issues to include education, gender equality for women and girls, water and sanitation issues. Working through partnerships with stakeholders at all levels is a key feature of UNICEF's approach.

28. The organization's governing body is the Executive Board, which provides intergovernmental support and oversight to the organization, in accordance with the overall policy guidance of the United Nations General Assembly and the Economic and Social Council (ECOSOC). The Board is made up of 36 Member States, who are elected by ECOSOC, according to a regional share of positions. The Board annually elects a President and a four-member Bureau that is responsible for bridging between the UNICEF secretariat and the regional groups represented on the Executive Board. The Bureau primarily deals with liaison, administrative and functional matters, such as facilitating the negotiation of decisions at the sessions, to enhance the effectiveness of the Executive Board. Members of the Bureau coordinate informal consultations within their respective regional groups.

29. The Board holds three sessions per year, and also participates in the Joint meeting of the Executive Boards of UNDP/UNFPA/UNOPS, UNICEF, UN-Women and WFP. The Bureau usually meets on a monthly basis, and daily during the Executive Board sessions.

30. Some of the recent key milestones in the life of the organization included:

- In 2010, the appointment of the current UNICEF Executive Director, who launched an organizational refocus on equity for children, accompanied by the development of the Monitoring Results for Equity System (MoRES).
- In 2013, the approval of the UNICEF Strategic Plan 2014-2017;
- In 2014, the Field Results Group was established, led by a new Deputy Executive Director post, to reinforce the focus on results and on performance monitoring in the field. While increasing the results orientation, this step also gave additional responsibilities to monitoring and evaluation staff in regional and country offices;
- In 2016, the UNICEF Policy on Research was endorsed.

31. Finally, in the upcoming Strategic Plan 2018-2021, UNICEF will align with the Agenda 2030's core concept, by committing to 'leaving no one, and no child, behind'.

### **3.2 Structure**

32. UNICEF headquarters are hosted in New York City, USA, where most of the corporate services and programme divisions are based. In addition, some central-level services and divisions are located in other countries, namely in Belgium, Denmark, Italy, Belgium, Japan and Switzerland.
33. UNICEF also comprises seven Regional Offices (ROs), listed here below with the area covered and the hosting country:
- CEE/CIS: Central and Eastern Europe and the Commonwealth of Independent States, Switzerland;
  - EAPRO: East Asia and the Pacific, Thailand;
  - ESARO: Eastern and Southern Africa, Kenya;
  - MENA: Middle East and North Africa, Jordan;
  - ROSA: South Asia, Nepal;
  - LACRO: Latin America and the Caribbean, Panama;
  - WCARO: West and Central Africa, Senegal.
34. Through the ROs, UNICEF reaches out to 159 countries, 117 of which host a Country Office led by an international Country Representative. Country offices implement UNICEF's mission through multi-year programmes of cooperation that are developed in collaboration with host governments.<sup>10</sup> The ROs and headquarters oversee this work and provide technical assistance to country offices as required.
35. UNICEF also has 34 National Committees. These are independent, local non-governmental organisations that serve as the public face of UNICEF in higher-income countries, by raising funds, engaging in advocacy and promoting UNICEF's visibility worldwide.

### **3.3 Resources**

36. UNICEF financial resources comprise:
- Regular Resources, or core financial resources, provided by Member States without any restriction on their use;
  - Other Resources (OR), which are funds from Member States and partners, earmarked for specific programmes and initiatives. Other resources include Thematic Funding, Other Resources Regular, Emergency Funding and Appeals, Pool Funding and Trust Funds.
37. Following a decision in 2011 by the Executive Board, UNICEF adopted an integrated resource plan for the presentation of budgetary information, harmonized for use by UNDP, UNFPA, UN-Women and UNICEF. The first UNICEF Integrated Budget was developed for the period 2014-2017 and proposed an institutional budget for the organization of US\$ 2,094.5 million, with US\$ 1,155.1 million provided through Regular Resources, US\$ 823.0 million in the form of cost recovery from Other Resources, and US\$ 116.4 million from Other Resources.

---

<sup>10</sup> Typically, UNICEF country programmes have a four to five years horizon, but in countries with protracted emergencies, country programmes tend to be much shorter, 1 to 2 years.

## 4 THE EVALUATION FUNCTION IN UNICEF: ORGANIZATIONAL AND FUNCTIONAL SET-UP AND RESOURCES

38. UNICEF evaluation function comprises an Evaluation Office (EO), staffed with evaluation advisers and specialists; and a few full-time, and many part-time evaluation specialists who respond, directly or indirectly, to directors of regional offices, representatives in country offices and to a few directors of headquarters divisions.
39. The Revised Evaluation Policy states that evaluation is a *shared* and *strongly decentralized* function, in line with the organization's structure. These two features, which pre-date the 2008 evaluation policy in UNICEF, represent its fundamental characteristics and inform and define all other relevant elements of the function, including the REP itself, the governance mechanisms, and various aspects of its implementation. Hence, it was considered that a description of the organizational set-up of the function, of its accountability framework at the different organizational levels, and of the resources available, would be useful to set the context for the analysis of all other aspects of the function.

### 4.1 The Evaluation Office

40. The Evaluation Office and its staff are based in UNICEF headquarters in New York. In January 2017, a new organisational chart placed the Evaluation Office as directly reporting to the Executive Director of UNICEF. Prior to this, the EO was located under the responsibility of the Deputy Executive Director (Management), together with Audit, Human Resources, Finance and Administration and Information Technology.<sup>11</sup>
41. In terms of mandate, the 2013 REP states that the Evaluation Office 'provides global leadership' for the function and defines its accountabilities in the areas of:
- a. governance and accountability, largely in a normative and monitoring role of the evaluation function;
  - b. conducting evaluations at the global level, including their identification and planning at the strategic level;
  - c. developing partnerships for evaluation, including supporting global networks for evaluations and collaboration with the international evaluation community;<sup>12</sup>
  - d. knowledge management on evaluations, including development of information systems for evaluation planning and monitoring, and tracking implementation of management responses; and
  - e. developing and professionalizing the evaluation function within the organization, including providing guidance on methods and approaches and on staffing of the function.
42. Among these, the EO staff are mostly engaged in commissioning and managing evaluations at the global level. The GEROS database indicates that in the period 2013-2016, EO conducted 24

---

<sup>11</sup> UNICEF High Level Structure, August 2015.

<sup>12</sup> This area also mentions 'especially by promoting national ownership and leadership of evaluation activities'; in the understanding of the Peer Review, this is difficult to imagine in practice, considering that EO is a fully headquarter-based office with no oversight role at regional nor country level.

evaluations, which represent 6.5% out of a total corporate output of 366 evaluations,<sup>13</sup> including: fifteen global evaluations; three Inter-Agency Humanitarian Evaluations and four evaluations of L3 humanitarian operations; and two joint evaluations with other entities of the UN system. In addition, the EO also produced two thematic synthesis and two meta-evaluations, and managed one evaluability assessment of the Strategic Plan.

43. Small teams of two or three staff members, led by a senior evaluation specialist,<sup>14</sup> have been set up for the different thematic areas. The standard steps in the evaluation management process typically include: liaising with the Programme Sections for preparing the Terms of Reference, launching competitive tender processes, selecting consulting firms, overseeing the work of evaluators' teams in headquarters and at regional and country level, quality assurance of both inception and draft evaluation reports, and sometimes of other deliverables too. All reports are approved by the EO Director before public release.
44. Two other important responsibilities of the EO, i.e. the monitoring of the performance of the evaluation function at corporate level and knowledge management on evaluations, are met by a small EO team led by a Senior Evaluation Specialist, that manages the following four key evaluation information systems:<sup>15</sup>
  - i. the Evaluation and Research Database (ERDB), publicly accessible, that consolidates all evaluation, research, studies and assessment report conducted in UNICEF;
  - ii. the Global Evaluation Reports Oversight System (GEROS), through which the quality assessment of final evaluation reports is made publicly available, as are the yearly meta-synthesis generated through the assessments;
  - iii. the Global Evaluation Dashboard, within the corporate dashboard, which reports on evaluation Key Performance Indicators at country, regional and global level; and
  - iv. the Evaluation Management Response Tracking System (EMR), wherein the Management Responses to each evaluation are available and progress on the implementation of each recommendation is recorded.
45. The four systems together largely cover management needs in terms of information on evaluation and provide the basis for the EO to report on the overall performance of the function. Another tool, VISION, allows monitoring disbursements for evaluations. A number of KPIs are attached to these systems, and are reflected in the Evaluation Dashboard, which is an important mechanism for visualizing performance.
46. An additional corporate system, the Plan for Research, Impact Monitoring and Evaluation (PRIME) is the electronic version of the Integrated Monitoring & Evaluation Planning (IMEP) system, and enables planning at the different organizational levels for the three evidence-generation activities. Strengths and weaknesses of these systems are discussed later in the report.
47. EO's performance in meeting its other responsibilities is discussed throughout the report, as these tend to be mainstreamed within the two main areas of work described above. It is nevertheless important to note here that due to the strong decentralization of the function, the role of the Office, in addition to conducting global evaluations, mostly consists of providing normative guidance, namely: issuing evaluation guidelines and templates; establishing and managing the

---

<sup>13</sup> Data down-loaded from GEROS on 3 April 2017. The Peer Review could not verify whether these figures reflect the entire universe of evaluations conducted in UNICEF.

<sup>14</sup> As of March 2017, only the Senior evaluation specialist for the Health sector was working, single-handed, on health-sector evaluations.

<sup>15</sup> At the time of the Peer Review, all UNICEF on-line systems were in the process of re-construction. Thus, the Peer Review looked at the services they provide and at links across them, but did not analyse any in depth from the view-points of functionality and user-friendliness.

quality assessment mechanism for evaluation reports; developing and managing the evaluation information databases; reporting on the performance of the function; and to a limited extent, internal and national evaluation capacity development. The EO has no responsibility for, nor possibility to make decisions on, evaluation activities conducted by the decentralized network.

## 4.2 The Decentralized Evaluation Function

48. The Decentralized Evaluation Function (DEF) includes all evaluations conducted under the responsibility of regional and country offices,<sup>16</sup> which represent the large majority of evaluations issued by UNICEF. For the period 2013-2016, the information in the GEROS database indicates that the DEF issued: 323 'country-focused' evaluations<sup>17</sup> (88.2%) addressing initiatives in 112 countries;<sup>18</sup> 16 evaluations (4.3%) managed and issued by five regional offices; and 3 evaluations managed and issued by programme divisions in headquarters. Forty-seven countries, or 29% of the total number of countries where UNICEF has some kind of activity according to the Country Office Reporting data-base, had no country-focused evaluation in this period. Also, two regional offices out of seven, did not manage nor issue any multi-country or regional evaluation in the same period. Box 1 below shows these data in matrix format.

**Box 1. Number of evaluations by country and region, in the period 2013-2016**

Region	N. country offices	N. countries	N. evaluations 2013-2016	Percentage of total by region, EO and programme divisions
<b>CEE/CIS countries + RO</b>	19	22	50	13,7%
<b>EAP countries + EAPRO</b>	14	27	39	10,7%
<b>ESA countries + ESARO</b>	21	21	75	20,5%
<b>EO</b>			24	6,6%
<b>HQ based Programme divisions</b>			3	0,8%
<b>LAC countries + LACRO</b>	21	37	41	11,2%
<b>MENA countries + MENARO</b>	15	20	24	6,6%
<b>SA countries + ROSA</b>	7	8	45	12,3%
<b>WCA countries + WCARO*</b>	20	24	65	17,8%
<b>Total</b>	<b>117</b>	<b>159</b>	<b>366</b>	<b>100,0%</b>

Source: GEROS and Country Office Reporting databases, elaborated by the Peer Review

49. In addition, UNICEF programmes, regional and country offices conduct many reviews, studies and research activities, which may be similar in scope to evaluations but do not follow the same process nor go through the GEROS final quality assessment. In 2014, an official corporate taxonomy for these different products was issued by the Office of Research and EO. Nevertheless, some confusion exists and in 2016, the preliminary screening process for quality assessment

<sup>16</sup> Evaluations directly or indirectly managed by headquarter-based divisions and sections are discussed in the next sub-section.

<sup>17</sup> The term 'country-focused' indicates evaluations that focus on work conducted in one single country, which may be a project, a theme, the whole country programme, the positioning of the organization in the country.

<sup>18</sup> See the Limitation section for an explanation on the data used.

through GEROS, identified three reviews that met all the evaluation criteria. Some interviewees mentioned that a number of evaluative efforts are classified as studies, reviews or research to bypass the evaluation-related procedures and quality assessment. This means that the information in GEROS could under-estimate the actual evaluative effort of the whole organization. The Peer Review considered that the issue would require more attention by UNICEF as a whole.

50. Typically, staff with evaluation responsibilities at country level prepare the annual Integrated Monitoring and Evaluation Plan (IMEP), the Costed Evaluation Plan (CEP) that is part of a Country Programme Document, develop terms of reference for the evaluations, select the evaluator/s through competitive tender processes or directly, support the conduct of the evaluations, revise and comment the inception and draft final reports, contribute to the dissemination of the final reports, prepare or contribute to the preparation of the Management Responses and update progress made in implementing the recommendations. All country offices should engage, at a minimum, in evaluation planning, every year for the IMEP and every four to five years for the Costed Evaluation Plan; the amount of time dedicated to this activity depends on the size and portfolio within the country programme. The evaluation-related work-load in each country office will depend on the number of countries for which the office is responsible, and on the number and type of evaluations that are carried out.
51. At the regional level, tasks include: planning for evaluations for the regional evaluation plans and support to country offices for national IMEPs and CPEs, real-time quality assurance to country level evaluations, or through Long Term Agreements (LTAs) with external consulting companies, commissioning and managing regional and multi-country evaluations, capacity development of country-level staff, participation in reference groups for country-level evaluations, networking with UN and non-UN regional evaluation networks. In one case, the Regional Evaluation Adviser could also be seconded to take part in UNDAF evaluations or multi-agency joint evaluations in the region.
52. An important feature of the DEF is the reporting line of staff responsible for evaluations and the accountabilities of heads of office in the process. The 2016 Self-Assessment indicated that only a minority, 23% of all respondents in the decentralized offices, report directly to the head of office, as follows:
  - Among the 12 respondents in the seven regional offices with the term 'evaluation' in their job titles: four Evaluation/M&E Advisers and Specialists report directly to the Regional Director, and oversee three other staff in total; and five, report to the Chief of Planning or Chief, M&E, who is at least one step removed from the head of the office;
  - Among 134 respondents with responsibilities for evaluation at country level: 30 (22%) report directly to the Country Representative; 48 (36%) report directly to the Deputy Country Representative, thus one step removed from the head of the office; and the other 56 (42%) report to a chief of section, or an M&E specialist, two or more steps removed from the head of the office;
  - Heads of offices approve the ToR and the selection of consultants in more than two thirds of evaluations; and the final report in more than half of evaluations; it can thus be assumed that in all other cases, staff at a lower level of seniority in the office, possibly more directly involved in programme management, would be responsible for approving these evaluation outputs.
53. No data was available from earlier reviews and assessments with respect to staff reporting lines that could be directly comparable with the data above.<sup>19</sup> The interviews with country-level M&E staff suggested that reporting to the Country Representative or to the Deputy Country

---

<sup>19</sup> The 2013 Review of the UNICEF Evaluation Policy and Function did analyse country M&E staff reporting lines, which indicated that 70% reported to the Country Representative or Deputy. This would suggest a step backward in this respect, as the Self-Assessment indicate that this situation applies only to 58% of the country offices. However, the two universes of reference may differ.

Representative is considered sufficient guarantee of independence while enabling a good knowledge of the programme and a good ownership of the evaluation process by the programme sections. However, interviewees also considered that reporting at a lower level could be problematic and generate conflicts.

54. Among Country Representatives, some considered that the M&E specialist reporting to a Planning/Monitoring/Evaluation Chief was appropriate, whereas others stated that they preferred the M&E specialist directly reporting to him/her, as evaluations should contribute to strategic decision-making at their level. A few within the latter group were in the process of making, or had just made, adjustments to the structure of their offices in this sense.
55. A particularly innovative model in this respect was launched in the Eastern Asia and Pacific region (EAPR), where a multi-country evaluation specialist position was agreed in 2014, became operational in 2015, and was recently extended for additional two years until 2019. An 'Employment Sharing Agreement' regulates the shared position at the P4-level among the country offices in Cambodia, Malaysia and Myanmar, with the purpose of "*strengthen their evaluation function, and its independence, support national ownership and capacity in evaluation. This shared position will also ensure that financial constraints will not affect the function*". In the view of one of the country offices, the arrangement is cost-effective by enabling access to a full-time evaluation specialist, at a reasonable cost for each of the concerned offices.
56. The incumbent, who is based in and reports administratively to the Country Representative in Cambodia, develops her work-plan in close consultation with two UNICEF Country Representatives, one Deputy Country Representative, and the Regional Evaluation Adviser; all provide inputs to the incumbent's performance assessment, too. The immediate result of the arrangement was an increase in the number and in the GEROS assessment of the quality of evaluations in the three countries.<sup>20</sup> The mechanism is also conducive to a stronger independence and credibility of the function, as the multiple reporting lines of the evaluation specialist diffuses any potential risk of stronger links with the programme units in each country office and/or of undue pressure from anyone in any of the countries; last, the position focuses exclusively on evaluation work.

### **4.3 Evaluations by headquarters divisions**

57. In addition to the EO and the DEF, a few headquarters divisions, responsible for both programmes and corporate functions, have established internal M&E units and are conducting their own evaluations, studies and reviews. This evaluation governance model is envisaged in the REP, which, among the accountabilities for evaluation, states that Division Directors "*are responsible for planning, resourcing and commissioning evaluations of the global policies and initiatives for which they are accountable, and for responding to relevant evaluation lessons and recommendations*".<sup>21</sup> Information in GEROS indicate that during the period 2013-2016, only three evaluations by headquarters divisions, excluding those by EO, went through the quality assessment process.
58. Different views were expressed by stakeholders in UNICEF headquarters on this model, which broadly corresponds to a country office evaluation function, only implemented by units that have a global or corporate-level mandate. Some interviewees explained that these units were established because, to some extent at least, global evaluations had not met the needs of their divisions in terms of relevance, quality or timeliness, or were not expected to do so in the medium-term.

---

<sup>20</sup> In Cambodia, one evaluation each year had been issued between 2013 and 2015, and two were completed in 2016; in Malaysia, only one evaluation had been conducted in 2015, and one was completed in 2016; in Myanmar a, only reviews had been carried out until 2016, when the country office completed four evaluations.

<sup>21</sup> Revised Evaluation Policy of UNICEF, E/ICEF/2013/14, April 2013, para 37.

59. The feedback through the interviews indicated that management of these evaluations is quite similar to the country level model, with M&E specialists reporting to a senior manager in the unit. The independence and credibility of the evaluations issued through this modality are likely to be the same as at country level. Similarly, it is unlikely these evaluations will meet strategic and priority needs for evaluative evidence at the corporate level, with the additional potential risk of duplication of the EO-managed evaluations.
60. Given the similarity between these evaluations and those conducted through the DEF, the Peer Review considered all as part of the same modality.

#### **4.4 Human resources**

61. In January 2016, the EO was structured as follows:<sup>22</sup> a Director at D2 level; six P5 level Senior Evaluation Specialists; four P4 and three P3 level Evaluation Specialists, one P3 Knowledge Management Specialist, one G7 and one G5 Administrative Assistants and one G5 Programme Assistant. One of the P3 positions was funded through the New Emerging Talents Initiative (NETI) and one additional P3 level evaluation specialist position was vacant. Some of the incumbent staff had been with the EO for several years and all P-staff already had professional experience in evaluation, or in Monitoring and Evaluation, when they had joined the EO. The gender balance in the Office was adequate.
62. By the time of the Peer Review in March 2017, the P5 level Senior Adviser Research and Evaluation, who was the most senior officer in the EO and played the role of deputy, had left his position in early 2017 in compliance with the corporate rotation policy. This is an important role, as under UNICEF rules, every office must have an official second in command to ensure business continuity and to have a complete line of managerial authority. Typically, the position requires both excellent management skills and in-depth technical knowledge in evaluation and excellent related competencies, as defined in the 2016 UNEG Evaluation Competency Framework.<sup>23</sup> The 2014/2017 EO Management Plan had proposed that a D-1 position be established for a Deputy Director in EO, though to no avail.
63. With respect to staff with responsibilities on evaluation within the DEF, the most recent data available indicate that in January 2017 UNICEF employed across 130 countries, 270 staff in country offices and 18 staff in Regional offices, whose job titles included the term 'evaluation'.<sup>24</sup> In December 2015, a comparable data set indicated that 66% of these positions were at P3 level or above, the required minimum level of experience for ensuring broad evaluation competence in UNICEF. At that time, the gender balance among staff with responsibility in evaluation was close to target at 48% for P3-level staff and above, but low, at 36%, for staff at P1-P2 levels.
64. Comparison with previous years showed an increase both in the number of country offices possessing a staff member with 'evaluation in the job title' and in the number of such staff in several country offices. The same report indicated that the percentage share of staff dealing exclusively with evaluation had remained stable over time at 9% of the total; whereas the share of staff combining only two different mandates, e.g. monitoring or planning, or research and evaluation, had decreased and reached 65% in 2015, while the percentage share of staff with three combined mandates had increased by 13% and reached 26%.<sup>25</sup>
65. Responses to the 2016 Self-Assessment, conducted by the EO, indicated that staff with responsibilities in evaluation at country level dedicate on average approximately 17% of their time

---

<sup>22</sup> EO organizational chart, February 2016.

<sup>23</sup> See <http://www.unevaluation.org/document/download/2610>.

<sup>24</sup> UNICEF ME Focal Points January 2017. This data-set did not include positions in UNICEF headquarters Programme Divisions.

<sup>25</sup> Annual report for 2015 on the evaluation function in UNICEF, E/ICEF/2016/11.

to evaluation. The amount of time dedicated to evaluation is larger in countries with larger programmes, reaching 21% for M&E staff working in countries with programmes above US\$ 50 million. At the regional level, responses to the same Self-Assessment indicated that the Evaluation/M&E Advisers and Specialists in the regional offices dedicated on average 39% of their time to evaluation, with only three out of 11 respondents in two regional offices, dedicating most of their time (85% or above) to evaluation. Some participants in the group discussions held in the context of the Self-Assessment felt they could not dedicate sufficient time to evaluation work and stated that they found it difficult *'to keep abreast of evolutions in the field'*, and that they scored their own competences in evaluation design and methodologies, as *'limited'*.

66. The data on staff-time spent on evaluation broadly explains the numbers of evaluations carried out in each country. It also indicates that, in addition to evaluation advisers and specialists in EO and in the regional offices, the total staff time in UNICEF spent on the DEF on a yearly basis by P3 level staff or above, corresponds to approximately 30 full-time staff. When compared with the evaluation functions of other similar UN organizations, this represents a high level of decentralized evaluation effort over seven regions. However, no analysis was ever done of what work-force would be required, to ensure an optimal evaluation coverage of UNICEF's work.

#### **4.5 Financial resources**

67. The 2008 evaluation policy established targets for evaluation budgets, in the range of 3-5% of programme resources to be allocated to evaluation, studies and research. The Revised Evaluation Policy led to a significant improvement in the overall funding mechanism by committing the organization to allocate *"a minimum of 1 per cent of its overall programme expenditure to evaluation"*.<sup>26</sup> This has become one of the KPIs for the evaluation function. At the same time, the EB requested UNICEF *"to allocate sufficient resources to the Evaluation Office in the institutional budget 2014-2017."*<sup>27</sup> This led to more than doubling (118% increase) the Regular Resources for the EO in 2014, which significantly reduced the Office's need to mobilize financial resources for conducting global evaluations.
68. UNICEF's Strategic Plan 2014-2017 (SP) clustered the EO and the Office of Internal Audit and Investigations under 'Independent corporate oversight and assurance' and allocated US\$ 7.3 million to the Evaluation Office, including a proposed increase of US\$ 6.7 million from the Regular Resources. This represented 0.3% of the proposed institutional budget of the organization for the SP that would be channelled to the Evaluation Office through the global and regional programme, for further strengthening and ensuring the sustained independence of the evaluation function.
69. As of March 2017, ten EO staff, including some professionals and three general service staff, were charged to the Corporate Human Resources Management. Five additional professional staff positions were funded by programme divisions, for managing global evaluations in Early Child Development (ECD), Education, Health, Water, Sanitation and Hygiene (WASH) and Humanitarian Operations.
70. The large majority of staff with responsibilities in the DEF were also charged to programme resources. By definition, all positions paid through programme resources are less stable; in the evaluation field, moreover, there is an added element of fragility due to the sensitivity of the function with respect to the required behavioural independence and the 'impartiality provisions'<sup>28</sup> needed to bolster this. This applies to staff in the DEF, and staff paid through programme, in the EO. Moreover, in an organization with significant decision-making authority at regional and country-level

---

<sup>26</sup> Annual report on the evaluation function and major evaluations, E/ICEF/2014/12.

<sup>27</sup> Ibid.

<sup>28</sup> UNEG Norm 5, Impartiality.

and a strong mobility policy, staff terms of reference are regularly subject to a certain degree of modification depending on the head of the office.

71. Since 2013 expenditures on evaluation have been captured through a corporate system called VISION, which tracks single disbursements against an evaluation code, and allows a very rapid visualization of resources spent on evaluation, aggregated up at country, regional and corporate levels, through the corporate Dashboard. This information is used for performance ratings at country and regional level on the budget KPI. However, interviews with country offices showed that; on the one hand, there is not a uniform understanding in the organization of what ‘evaluation costs’ include and what type of disbursements should be charged to the evaluation codes; on the other hand, human errors do occur in the coding of single activities, which cannot be later modified. This means that there are likely to be inconsistencies in the coding system, considering the large number of annual entries for evaluations, and the possible different interpretations of what qualifies as evaluation expenditure across the many offices responsible for entering the data. This suggests that the overall information generated at country, regional and corporate may not be fully reliable in the absence of both clear guidelines of what should be classified as evaluation expenditure and of ex-post checks and adjustments.

72. The system is useful nevertheless, and has shown a progressive increase in resources spent on evaluations, albeit still short of the target of 1%. In 2016, corporate expenditure on evaluation was US\$ 35.5 million, corresponding to 0.7% of the Total Budget Use, including the expenditure by the EO. Given the huge diversity in the financial size of country programmes and single interventions, the resources dedicated by every single office to each evaluation do vary significantly.

73. For example, Box 2 below shows a highly-disaggregated snapshot of the cost of evaluations completed in 2016, against the financial resources allocated to each of the evaluated programmes in Cambodia, Malaysia and Myanmar.

**Box 2. Cost of evaluations compared to the budget of evaluated interventions, 2016**

N.	Country	Budget of the intervention (US\$)	Evaluation cost (US\$)	Evaluation cost as % of intervention budget	Type of contract
1	<b>Cambodia</b>	20,000,000	71,500	0.4%	Institutional Contract
2	<b>Cambodia</b>	733,091	44,740	6.1%	Individual Contract
3	<b>Malaysia</b>	800,000	76,572	9.6%	Institutional Contract
4	<b>Myanmar</b>	4,111,724*	159,209	3.9%	Institutional Contract
5	<b>Myanmar</b>	215,367	95,090	44.2%	Institutional Contract
6	<b>Myanmar</b>	108,178	98,146	90.7%	Institutional Contract (LTA)
7	<b>Myanmar</b>	509,085	107,979	21.2%	Institutional Contract (LTA)

Source: Evaluation Specialist (Cambodia, Malaysia and Myanmar)

\*: personnel costs included in the budget.

74. Evaluation costs were variable, though most broadly within the same order of magnitude. What varied dramatically was the ratio between the cost of the evaluation and the budget of the interventions, even keeping in mind that the real budgets of the interventions were higher, because the budgets in the column ‘Budget for the intervention’, did not include personnel costs, with the noted exception of N<sup>o</sup>.4.

75. In the case of evaluations N<sup>o</sup>. 6 and 7, donors requested that evaluations be carried out to assess the effectiveness of the approach before considering the possibility of scaling it up. The two evaluations, which were positively rated through GEROS as ‘highly satisfactory’ and ‘satisfactory’ respectively, eventually confirmed the overall validity of the approach whilst highlighting the

necessary adjustments. Most recommendations for both were agreed upon and were either completed or in progress six months after the upload of the Management Response in the tracking system. This was a clear case where the specific judgment at country level on the utility of an evaluation was required and resources well above 1 % had to be made available, as foreseen by the REP with regards to evaluations of innovative programmes before replication or scaling-up.

76. There is no doubt that UNICEF Senior Management commitment to achieving the target of 1% of programme resources is strong; the budget target for evaluation is well known across the whole organization and provides visibility and importance to the evaluation function by attracting Management's attention. This has also been a recurrent topic in the Annual reports on the Evaluation function to the Executive Board, which has systematically recommended improvements in this respect.
77. However, the blanket application of any target at all levels and for all units may not be the most efficient use of resources, and a frequent comment made to the Peer Review was that the rule should be applied with flexibility and be 'inspirational' rather than prescriptive. A very large country programme would be likely to receive an over-allocation of evaluation resources, whilst the amount corresponding to 1% in shrinking country programmes in Middle-Income Countries would allow financing of, at best, one evaluation over a five-year period.
78. In addition, the Member States in which UNICEF, and other UN organizations work, are changing rapidly. On the one hand, the work of the UN is increasingly moving more upstream, towards support on national and sectoral policies and strategies. Although there are still several countries where this shift is not taking place, in general the trend is towards smaller country offices and smaller presence on the ground, with jointly implemented programmes with national governments. At the same time, the Sustainable Development Goals and Agenda 2030 firmly put the need for learning and accountability at the national level. The UN evaluation system will have to adjust to these trends, making sufficient resources available where these will be more efficient and useful to ensure sufficient evaluative coverage, capacity development and accountability. A variety of disbursement modalities will likely be required, including joint evaluations at country level and evaluations of regional or multi-country programmes.

## 5 MAIN AREAS OF ANALYSIS

79. This section analyses and discusses each of the seven areas identified in the Terms of Reference for the Peer Review. Although an earlier draft of the report was structured by, and discussed issues one by one, this final version presents the findings in one single narrative for each area, to keep repetitions to the minimum.

### 5.1 Evaluation Policy

80. The Evaluation function in UNICEF has existed for many years. In 2002, the first strategic plan for evaluation in UNICEF<sup>29</sup> had been presented by the Director of the Evaluation Office to the Executive Board and had “*situated the role of evaluation among the instruments of oversight and as a tool that supports programme excellence...*”. At the time, the function was already fully decentralized and evaluation reports had just been made public on the Internet, which was appreciated by the Executive Board members. The latter also requested: a more systematic use of external evaluators in conducting evaluations; that programme countries be supported to evaluate their own programmes; and that ‘evaluations of relevance’ be brought to the Board’s attention.

81. In 2006, the first Peer Review of the UNICEF evaluation function was conducted, to assess the function against the core pillars of independence, credibility and utility. It concluded that UNICEF needed a single evaluation policy, consistent with the recently approved UNEG Norms and Standards (N&S), which would assert the independence of the function through a direct reporting line to UNICEF Executive Director. It also stated that the policy should address: the contribution of evaluation to learning, accountability and decision-making; the strong decentralized structure of the organization; and clarify roles and responsibilities and accountabilities at the central, regional and country level. The Executive Board at its June 2006 session welcomed the Peer Review report and requested that UNICEF prepare a comprehensive evaluation policy to be considered by the Board.

82. In 2008, the Board approved the first UNICEF Evaluation Policy and resolved that the Board itself would exercise oversight of the function. The policy largely met UNEG N&S; it defined fifteen principles for evaluation, including on the independence and credibility of the function, roles and responsibilities for the Evaluation Office and for evaluation planning, resource allocation, management and quality assurance at headquarters, regional and country level. At the same time, the policy distanced evaluation from the oversight function and located it closer to research and to the evidence-generating function.

83. In 2009, UNICEF Executive Director signed an Executive Directive, to provide operational guidance for the implementation of the Evaluation Policy. Accountabilities for each level within the organization were detailed, and stronger links between the various levels in the function established, with the EO responsible for revising the regional evaluation plans, clearing recruitment of M&E officers, providing strategic guidance to the other levels on various aspects of the policy, among other responsibilities. To some extent, the Directive counter-balanced the Policy by emphasising the internal coherence and independence of the function.

84. Over the following years, the Executive Board’s decisions on the evaluation function recurrently stressed the independence of the function and the importance of a number of issues, including: evaluation capacity development at country-level, both within the organization and at the national level; monitoring of evaluation activities; integration of cross-cutting areas in evaluation methodologies, including in human rights-based approaches and gender analysis; compliance with

---

<sup>29</sup> Report on the Evaluation function in the context of the Medium Term Strategic Plan, E/ICEF/2002/10, [https://www.unicef.org/about/execboard/files/2002-8\\_Rev.1\\_Report-ODS-English.pdf](https://www.unicef.org/about/execboard/files/2002-8_Rev.1_Report-ODS-English.pdf).

the management response process; and integration of evaluation results into policies and programmes.

85. In 2012, the Board requested UNICEF to undertake a review of the evaluation function and policy, as foreseen in the Evaluation Policy itself. The review, completed in early 2013, identified several key issues and formulated three major recommendations for the Policy, and five for its implementation. The issues tackled included: a target for financial resources exclusively allocated to evaluation; the need for more clarity on the role of Monitoring and Evaluation specialists at country level; the need to strengthen the EO's role in providing guidance and oversight on the evaluation process across the organization, including on independence and transparency of the evaluation process; more efficiency required in UNICEF's work in support of national evaluation capacity development; and evaluation coverage. In addition, a revision of the Evaluation Policy was recommended.
86. The 2013 Revised Evaluation Policy (REP) was developed accordingly and the Executive Board endorsed it, with the request to UNICEF to: allocate sufficient resources to the evaluation function; maintain close consultation with national governments in evaluations at country level, including the use of national evaluation systems; and integrate evaluation results in policies and programmes.
87. The basic tenets of the REP, that inform the theory and practice of evaluation in UNICEF, are:
- i. the accountability framework, that integrates the evaluation function into the highly-decentralized structure of the organization;
  - ii. the integration of the function in the mandate and policies of the organization;
  - iii. the attribution of the responsibility for safeguarding both financial and staffing resources for evaluation to the Executive Director;
  - iv. a strong emphasis on the importance of the use / utility and intentionality of evaluations; and
  - v. the contribution of evaluation to the evidence-generation function.
88. Each of these elements is discussed here below. Similar to the 2008 Evaluation Policy, the REP was also informed by the UNEG N&S; the Peer Review thus analysed it against the updated N&S approved by UNEG in 2016, with a view to identify areas and elements where adjustments might be desirable.<sup>30</sup>
89. With regards to the accountability framework, the REP specifies who should do what at the different organizational levels, with a reasonable level of detail, in terms of planning, commissioning, managing and using evaluations at each level. The measures foreseen for independence and impartiality entail the exclusion of staff from designing and managing evaluations of programmes, for which they held 'direct responsibility'; and from being a member 'strictu-sensu' of an evaluation team. The Policy also refers to the need to establish appropriate arrangements to avoid conflict of interest "*to support the impartiality and independence of the evaluation function by separating programme management responsibilities from evaluation duties*".
90. These are fundamental measures, which however, in the strongly decentralized structure and culture of the organization, do not seem sufficient to ensure the separation of duties and responsibilities between planning and management on the one hand, and evaluation on the other, which should be the cornerstone for evaluation to be independent and credible. Thus, the REP accountability framework appears weak with respect to setting out the safeguards for the overall independence, organizational and behavioural, of staff responsible for managing and conducting evaluations, and for the impartiality of the function, against both 2005 and 2016 versions of the UNEG N&S.

---

<sup>30</sup> For a full comparison of the REP against the 2016 UNEG Norms and Standards, see Annex 6.

91. Typically, UNICEF managers stated that the current accountability framework where heads of units and offices are responsible for the function is a guarantee per se that evaluations will be carried out in an independent manner and all Regional Directors accepted that the evaluation function was their responsibility. This suggests, as confirmed through interviews with other categories of stakeholders that UNICEF as an organisation tends to understand evaluation to be a manager's programmatic tool, close to research and Results-Based Management (RBM), rather than a distinct and independent function and tool that should be useful for management and programmes, without being controlled by them.
92. In practice, the oversight and accountability element of evaluation are often reduced to a minimum. It is moreover debatable that assigning responsibility for its oversight and management to highly competent managers in other areas, with over-busy agendas, is an efficient and effective manner to manage a function that should also contribute to accountability and oversight.
93. The 2016 Self-Assessment showed that the independence of the function and the assignment of accountabilities within it linked to the level of its decentralization, were areas of main concern for respondents. The Peer Review was indeed told of a few situations suggesting undue influence on staff managing evaluations, at different levels. Actions to manage the conflict were not taken in all cases, nor to prevent this from happening again. Furthermore, the contractual situation of staff paid from programme resources, can indirectly lead to self-censorship and Evaluation/M&E Advisers and Specialists at regional and country level do not have any institutional channel to revert to, should they perceive undue pressure on the independence and impartiality of their evaluation-related work.
94. The Self-Assessment also showed concerns about the 'threat' represented by the mandatory corporate rotation policy for the professionalization of the evaluation function and for supporting behavioural independence of evaluation staff. Arguably, an evaluation policy should not run counter to other policies in any organization. Still, the specific requirements for professional competence and behavioural independence of staff with evaluation responsibilities would require that an evaluation policy clearly refers to: the professional competence of evaluation staff, the career safeguards, and the criteria for rotation to and from the evaluation positions. Different options could be taken into consideration that would enable the independence and credibility of the evaluation function in UNICEF to be enhanced. These include: the adoption of staffing criteria for evaluation more closely aligned to those applied to staff recruited for the Audit function; the rotation of evaluation staff only across evaluation positions at headquarters, regional and country level; a mix of rotational and non-rotational positions in headquarters, as is the case in the WFP evaluation function.
95. Overall, the REP accountability framework is not fully aligned with international evaluation standards that consider the segregation of roles and responsibilities a basic pillar of the evaluation function. This is affecting both the impartiality and credibility of the entire evaluation process and its products.<sup>31</sup> This weakness had already been identified in the 2014 analysis by the Joint Inspection Unit of the United Nations (JIU), that suggested that UNICEF should re-examine its policy "*for the structural independence of the evaluation function and decide on how best to enhance and safeguard structural independence, including in the appointment of the head of the evaluation unit*".
96. With respect to the 'functional and organisational arrangements', the REP foresees that "*UNICEF will develop a comprehensive evaluation strategy and plan of action to carry forward effective implementation of the Evaluation Policy*", to be complemented by regional strategies. Although two Regional Offices had developed, or were in the process of developing, their respective regional evaluation strategies by the time of the Peer Review, no such tool had yet been

---

<sup>31</sup> See in particular Norms 4 and 5, 2016 UNEG Norms and Standards.

prepared for the REP.<sup>32</sup> Four years into the implementation of the REP, an agreed and well documented mechanism through which the decentralized evaluation activities could better contribute to the coherence, effectiveness and usefulness of the evaluation function for the organization, above the local level, is still pending.

97. The 2016 Self-Assessment showed that staff perceives the need for more clarity and guidance on the implementation of the REP, including on reporting. Among the Issues for Consideration, it was proposed that *“Key policy statements should be accompanied by executive directives and certain elements could be broken down and operationalized in an action plan”*.
98. With respect to the integration of evaluation into the policies of the organization, the REP refers in clear terms to the role of the evaluation function in support of UNICEF’s mandate, firmly aligns the evaluation function with UNICEF strategies, and links evaluation planning to the organization’s strategic objectives. Furthermore, the UNICEF Strategic Plan 2014-2017 refers to Evaluation as a tool to assess programme results and as one of the pillars for achieving the objectives of the SP, complementing other evidence-generating functions. The SP also reaffirms the emphasis given through the first Evaluation Policy to the knowledge-generation and lessons learning purpose of evaluation, over the accountability function.
99. The REP assigns the ultimate responsibility for the evaluation function and for its resources, to the Executive Director. It also recognizes that high-quality evaluations require adequate financial resources and competent human resources; and states that the establishment of evaluation accountabilities in each office will require some in-house evaluation capacity at all levels, complemented by external skills and competences, catered for through consultants, creative arrangements across UNICEF offices, and reliance on regional or national centres of excellence. The EO was assigned responsibility for internal capacity development, through training and technical support. In contrast to the provision for financial resources, arrangements for human resources are largely left to the initiative of individual offices. Thus, overall, the REP adequately addresses the provision of human and financial resources for evaluation, in a coherent manner with the overarching principle of a highly-decentralized evaluation function. However, since the REP gives the EO no oversight nor functional link with M&E specialists beyond the Office itself, this task appears difficult to accomplish.
100. In the same vein, and as foreseen by the UNEG N&S, the independence of the Director of the Evaluation Office is ensured through a reporting line to the Executive Director of UNICEF. However, the policy introduces a *“day-to-day supervision by the Deputy Executive Director (Management)”* which should be removed in a future policy revision as it might create perceptions that dilute the independence of the function.<sup>33</sup> As already mentioned, the Policy is also silent on safeguards for the independence of staff working on evaluations.<sup>34</sup>
101. The revised N&S also specify in detail the required competences for the heads and staff of UN evaluation units. For the former, these include technical and professional skills in evaluation, *“including a stronger knowledge base on evaluation to enable providing substantive guidance on global issues and evaluation trends”*, mastery of evaluation ethics in complex contexts; management skills; and communication and interpersonal skills. The REP does not specify anything in this respect, apart from a generic statement that high-quality evaluations *“possess core evaluation competencies”* and relevant experience.

---

<sup>32</sup> The development of a comprehensive global evaluation action plan to support the implementation of the REP had been agreed with the Office of Research as a priority for 2015 and 2016, and had been included in the Annual Work-Plan for the EO in 2016.

<sup>33</sup> Until 2017, this practice was also reflected in the 2015 organizational chart, which missed the direct link foreseen by the Policy between the EO Director and the Executive Director.

<sup>34</sup> When revising the evaluation policy, UNICEF may wish to consider the relevance of the ‘impartiality provisions’, articulated in the WFP evaluation policy, and develop something along similar lines, tailored to the UNICEF context.

102. Furthermore, the REP, by stating that evaluation is a shared function that requires the contribution of many professional staff, and by assigning accountabilities for the function at the decentralized level exclusively to regional directors and country representatives, yet without clearly specifying competences, skills, and arrangements for impartiality in managing evaluations, runs contrary to the UNEG N&S concepts for professionalism in evaluation. Applying to both evaluators and evaluation managers the N&S state that *“the evaluation function is staffed by professionals with evaluation competencies in line with the UNEG Competency Framework.”*
103. With respect to evaluation contributing to the corporate evidence-generation function, the REP states that Evaluation and Research are closely related and that both contribute to the knowledge agenda of UNICEF, although from different perspectives: Research seeks theoretical knowledge, while Evaluation tests the achievement, relevance and sustainability of results. This approach is also confirmed in the UNICEF Policy on Research, issued in 2016 and which defines evaluation, research and data collection, as the three key pillars of UNICEF's evidence-generation efforts.
104. However, the REP does not foresee corporate mechanisms for evaluation use and influence, in addition to the management response process and a tracking system for recommendations implementation. This leaves a conceptual and practical gap between the use of evaluation results at the local level, and the scaling up of findings and lessons learned at a broader corporate scale. In other words, UNICEF lacks a proper Theory of Change for its evaluation function; in addition to clarifying the overall contribution of the function to the organization, this would help in driving adjustments in the evaluation process at all levels and in all steps.
105. With respect to audit and monitoring, the REP states that evaluation is distinct from these functions. In this, the policy is in line with UNEG N&S as well, although there is a certain level of internal contradiction when it refers to ‘monitoring and evaluation specialists’ responsible for managing evaluations at country level through the whole process. In practice, outside the EO, the large majority of evaluations in UNICEF are managed by staff who typically dedicate most of their time to other functions, including research, monitoring, etc., thereby blurring the distinction between these roles and those of the evaluation function.
106. Finally, the Peer Review also considered whether other external contextual changes had occurred, with a bearing on UN evaluation policies in general. The extent to which a new UNICEF Evaluation Policy should, or might wish to, take these into consideration is discussed here below.
- i. In 2014, the United Nations General Assembly approved Resolution 69/237 on “Building capacity for the evaluation of development activities at the country level”, calling on the UN for supporting Member States in developing their national evaluation capacity. Given the long-term experience of UNICEF in National Evaluation Capacity Development, the REP had already dedicated a well-developed section to it and made provisions for UNICEF's engagement in NECD. There seems to be no need for any adjustment to the policy per se in this respect, apart from, possibly, an explicit reference to the Resolution.
  - ii. In 2015, the UN General Assembly approved Agenda 2030 and the Sustainable Development Goals, which provide a new framework of action for Member States and for the whole UN system for fifteen years. The UNICEF new Strategic Plan 2018-2021 appears to be fully informed by the concept at the core of the SDGs, by aiming at ‘leaving no one, and no child, behind’. This is a strong call for equity, which UNICEF evaluation function had already fully embraced in its conceptual framework by adequately reflecting it in the REP. Nevertheless, Agenda 2030 also poses different challenges to the UN evaluation functions. UNEG has already carried out some work in this respect and more is likely to be conducted in future; although no formal decisions have been made so far in this respect by UNEG Heads, results from these analysis and discussions might become an important reference to be taken into consideration, in any future revision of UNICEF evaluation policy.

- iii. In late 2016, the UN General Assembly approved a new resolution on the Quadrennial Comprehensive Policy Review of the UN. Among others, the UN system is called to support Member States in their evaluation capacity, and to promote more integrated support including in evaluation. The resolution also underscores the importance of an independent and impartial system-wide evaluation mechanism “*that enhances coherence and interdependence in the overall evaluation architecture of the United Nations development system*”. Similar to the approval of Agenda 2030, this resolution will require adjustments in the UN evaluation system that UNICEF may wish to take into consideration, including at the policy level.
107. Last, the REP does not appear to be sufficiently clear with respect to integrating a human rights perspective into evaluations, and to the diffusion of evaluations. These are minor issues, considering the strong commitment of the UNICEF evaluation function to equity-based evaluation and to the use of evaluations; nevertheless, an explicit reference to both in a future version of the policy, would be appropriate.

## 5.2 Governance arrangements

108. In 2008, upon approval of the first Evaluation Policy of UNICEF, the Executive Board had resolved that oversight of the evaluation function would rest with the Board itself. This oversight role remained in the 2013 REP, which however firmly assigned the ‘*safeguard of the integrity of the evaluation function and its independence*’ to the Executive Director of UNICEF, to whom the Director of the EO, as well as Division and Regional Directors and Country Representatives, are accountable.
109. Over the years, the EB has raised recurrent requests to the organization, on: adequate financial resources allocated to evaluation; better coverage of the organization’s work through evaluation; better use of evaluations, through management responses, their follow-up and integration of evaluation findings into policies and programmes; strengthening of the decentralized evaluation function. In general, trust, interest and support seem to inform the relationship among the Executive Board, the EO and the corporate evaluation function at large. The EB’s decisions on evaluation have addressed over time several aspects of the function; in a number of cases, the EB formulated ‘requests’, which require a follow-up.
110. Evaluation is a standing agenda item in three of the Board’s sessions every year, under the heading ‘Evaluation, audit and oversight matters’.<sup>35</sup> In the June session, the EO presents to the Executive Board the Annual Report on the Evaluation Function in UNICEF, which discusses the performance of the evaluation function at the corporate level, beyond the direct responsibility and mandate of the EO itself. The reports provide information on all the established KPIs, as well as on the human and financial resources for EO, and typically discuss one topic in depth. Until 2013, the reports also provided an analysis of findings from major evaluations, but the Executive Board asked the latter information to be made available through another, separate report. Since 2014 Management has prepared a response to the Annual Report on the Evaluation Function to comply with a specific Board request.
111. In recent years, these reports have been candid assessments of progress, strengths and weaknesses of the function at the different levels. The EB showed appreciation for this approach and the reports of debates over evaluations suggest lively discussions and well-informed participants, and appear to be conducive to the use of evidence in decisions made by the Executive Board itself and by Senior Management.

---

<sup>35</sup> The only Board’s meeting in any given year where evaluation is not discussed is the Joint Meeting of the Executive Boards of UNDP/UNFPA/UNOPS, UNICEF, UN-Women and WFP.

112. The only shortcoming in the Annual report has been the timeliness of information on the KPIs on evaluation coverage and quality, due to the lack of synchronization between deadlines established for uploading evaluation reports in to the ERDB and the presentation of the Annual report to the Board.<sup>36</sup> This lag-time in reporting may have reduced the sense of urgency for the accountability feedback provided by evaluations, as well as the attention to deliberations by the Executive Board which come 'late', by definition, with respect to the actual context. However, the time-lag problem has recently been resolved with the Annual report for 2016 on the evaluation function in UNICEF, presented to the Executive Board at its 2017 Annual Session. This undoubtedly is a key step forward toward stronger accountability and usefulness of the function.
113. Other evaluation agenda items presented to the Board have included; the 2006 Peer Review report and the 2012 review of the function, the evaluation policies, and the results of the 2016 Self-Assessment. The Evaluation Office also presents to the Executive Board, reports of global evaluations conducted by the EO itself, together with the respective Management Responses. However, this key window of opportunity for visibility and influence appears not to be utilized to its full potential, as the selection of which evaluations to present is at the discretion of the EO Director, reportedly in consultation with Senior Management.<sup>37</sup> In the absence of official standard criteria and a well-defined process for making the choice, this step risks being perceived as ad-hoc and not fully transparent, as also noted by the latest Multilateral Organisation Performance Assessment Network (MOPAN) report. In the view of the Peer Review, the systematic presentation of all global evaluation products, and if time allowed also of important evaluation reports delivered by the DEF, would help promote the use of evaluation findings and enhance the transparency of the evaluation function.
114. The Evaluation Office is accountable to the ED through line management structure, and reports to the Executive Board. The REP specifies that the EO Director is appointed by the Executive Director, and that the incumbent is entitled two terms of four years each, with no re-entry into the organization at the end of the second term. Since 2010, the position of Director of the Evaluation Office has been funded at the D2 level, in line with UNEG N&S. However, the REP does not make any provision on specific professional competences, for example in the form of minimum requirements for the technical competence and experience in the evaluation field of the appointee, nor in terms of avoiding potential conflicts of interest when UNICEF internal candidates from outside the evaluation function apply for the position. In this respect, the REP misses the opportunity to make a specific reference to UNEG Competencies for Heads of evaluation functions or offices and to ensure that external evaluation professionals are members of the selection panel, a practice adopted by UN agencies of similar size and with evaluation directors at the same level of seniority. No specific arrangements have been developed either for performance management or termination of the appointment, which is subject to standard UNICEF human resources rules for senior managers.
115. The EO reports to Management on a frequent and regular basis, as foreseen by the REP, mostly through the Deputy Executive Director (Management), a satisfactory arrangement in the view of Senior Management. In the view of the Peer Review, regular engagement and dialogue with senior management is to be encouraged but such should not be construed as a formal reporting line (other than to the ED). The EO also prepares an annual report on its own specific work, which goes to Management only and is available through the UNICEF intranet.
116. UNICEF also established an internal Global Evaluation Committee, whose ToR were last revised in 2009. The Committee is chaired by the Executive Director and comprises '*the Deputy*

---

<sup>36</sup> Reports for the Executive Board have to be prepared well in advance the actual meeting, for editing and translation purposes.

<sup>37</sup> In the period 2013-2016, out of the 24 global evaluation reports and syntheses issued by the EO, records indicate that 16 were presented to the Executive Board.

*Executive Directors, Directors of Programmes and of Policy and Practice*. As of 2017, the Regional Directors were also members. The EO Director attends all meetings, and the Evaluation Office acts as secretariat to the Committee. The ToR also mention one external evaluation specialist, though the Peer Review had no evidence of anyone in this role. The purpose of the Evaluation Committee is to engage Senior Management, in an advisory role, with the work of the evaluation function. Its tasks include the review of:

- The Evaluation Office work programme and budget for evaluation;
- UNICEF evaluation reports that have relevance at the global governance level;
- The annual report to the Executive Board on the evaluation function; and
- Management response processes and follow-up to evaluations.

117. Members of the Global Evaluation Committee asserted its usefulness, mostly as a sounding board for new ideas and proposals on evaluation matters. Recently, the Committee has not been very active and convened only once or twice per year, against the foreseen three or four meetings in its ToR. In the view of the Peer Review, the Committee can fulfil an important role in supporting various aspects of the evaluation function, from the planning of corporate-level evaluations, to discussing mechanisms and commitments to financing of the evaluation function, and to diffusion and use of evaluation results. In this respect, in compliance with its ToR, it should hold its meetings more frequently, and include an external evaluation adviser to provide an independent view and knowledge about evaluation matters.

118. The Director of the Evaluation Office participates in the Management Team meetings in headquarters and, annually, in Regional Management Team meetings where he has contacts with Regional Directors and Country Representatives. These are intended to be occasions for sharing information and advocating for evaluation, not for oversight. Overall, the EO is perceived as effectively contributing to corporate initiatives where their knowledge and know-how is particularly relevant. Two examples mentioned by interviewees were the collaboration with the Office of Research to draft the UNICEF Policy on Research, inspired by the REP; and the Electronic Content Management (ECM), a five-year project that will replace the entire UNICEF's intranet. On the latter, EO provided oversight of the financial aspects of project implementation and brought in an important user's perspective, that helped shaping the performance measures for the ECM. Taking into account the corporate accountability framework for the function, participation of the EO Director and of EO staff in these meetings and efforts, did not appear to create any conflict of interest in the understanding of the Peer Review.

119. As already discussed, the oversight of the decentralized evaluation function is decentralized to each organizational level. The Evaluation Office only has an advisory and advocacy role for the DEF, in addition to its monitoring and reporting duties on performance, through the Annual Report on the Evaluation Function for the Executive Board, based on the different KPIs that are consolidated in the corporate Evaluation Dashboard. Regional Directors and Country Representatives stated they were quite attentive to it. The EO also maintains regular contact through virtual meetings with the network of Regional Evaluation/M&E Advisers and Specialists, mostly used for information sharing. However, links between the EO and M&E Specialists in the country offices are limited.

120. The REP proposed a Peer Review of the evaluation function to be carried out in 2015, two years after its approval. This might have been somewhat early for the REP to have proved its worth. The present exercise took place four years later, which also coincided with a change in leadership in EO. Thus, despite the significant delay in launching it, the Peer Review was still conducted at a useful and appropriate moment in time.

121. The performance of the evaluation function is also assessed by the Audit Advisory Committee (AAC), which is an external and independent body that meets several times per year and analyses various aspects of UNICEF institutional set-up and operations. AAC reports are publicly available; in recent years, evaluation issues have been frequently addressed, including discussions on the independence of the function and on the implementation of the electronic Integrated Monitoring and Evaluation Plan (IMEP). The AAC also formulated recommendations, namely on the length of term of the Director of EO and on the improvement of the timeliness and quality of management responses to evaluations.
122. Every three to four years, another regular assessment comes in through the MOPAN, which systematically analyses the organisation, including the evaluation function. The latest MOPAN was conducted in 2016 and the report issued in early 2017. The report stated that *'the quality of evaluations is improving overall, supported by robust quality assurance systems'*, though it also noted that evaluation coverage *"with emphasis on robust evaluative evidence on the efficiency and sustainability of UNICEF programming"* was the evaluation-related area where improvements were identified as desirable.

### **5.3 Evaluation planning and coverage**

123. In compliance with the REP and its accountability framework, each office in UNICEF, from the EO to country offices, is responsible for planning evaluations at its level. Regional Evaluation/M&E Advisers and Specialists contribute to and review regional and country-level evaluation plans; whereas the Evaluation Office is responsible for corporate evaluations and no longer has the oversight responsibility, the 2009 Executive Directive gave to it, over the regional evaluation plans.
124. Country offices, and possibly regional offices and programme divisions as well, used to develop paper-based annual plans on evaluation, research, studies and M&E capacity-building through the corporate Integrated Monitoring & Evaluation Planning (IMEP) system. In 2015, the Plan for Research, Impact Monitoring and Evaluation (PRIME) was launched, a result of an EO-led effort with extensive inputs from UNICEF M&E stakeholders and the Office of Research. PRIME is the updated electronic version of IMEP, which enables more detailed planning of single activities at the different organisational levels for the three evidence-generation activities.<sup>38</sup>
125. In 2014, the EO developed a four-year Global Evaluation Plan, comprising a list of 'major' and 'smaller' evaluations, as well as synthesis studies, to be carried out over the following four years. 'Major' evaluations aim at providing evaluation evidence on the seven outcome areas and seven implementation strategies set out in the Strategic Plan. Principles for their identification were: contribution to a coherent framework of the organization's results, coverage, interest, demand, relevance, risk, and additionality (no duplication). The EO planned to deliver, on an annual basis, four major evaluations, two smaller evaluations (corporate strategies), two evaluation syntheses and one methodological study. The Plan also gives due attention to other areas of work, including National Evaluation Capacity Development, Knowledge Management, methodology development, quality assurance, inter-agency collaboration, which all together cover all the responsibilities assigned to EO in the 2013 REP; and it foresaw a mid-term review of the Plan itself.
126. The Plan did not include evaluations of humanitarian operations, based on the stated 'unpredictability' of such work. Although this is perfectly reasonable with respect to natural disaster-related operations, e.g. Typhoon Haiyan in the Philippines or the Ebola epidemics in West Africa, it could be argued that at the time of developing the Plan, evaluations of protracted crises in which

---

<sup>38</sup> Information on IMEP and PRIME was extracted April 2017, from the PRIME Web site at <https://intranet.unicef.org/epp/evalsite.nsf/8e1ddc662803020785256ede00706595/48c3a665a6a9ba3a85257ede00706052?OpenDocument>.

UNICEF was participating with humanitarian operations, in the form of country-programme evaluations, could have been factored into the Plan itself as they were likely to happen.

127. The Plan did not explain in detail the extent of consultation with the rest of the organization for its preparation. Reference is made to other documentary sources, e.g. audit reports and reviews by the Joint Inspection Unit (JIU) and bilateral entities. Feedback from interviews indicated that EO consulted to some extent with Senior Management, including through the Global Evaluation Committee; and that consultations do take place in the case of evaluations of emergency operations. Nevertheless, various stakeholders noted in the interviews, that room exists for improving consultation on the planning of global evaluations.
128. Also in 2014, the Executive Board asked UNICEF to prepare Costed Evaluation Plans (CEP) as a standing element of new Country Programme Documents submitted to the Board for its approval. A CEP consists of a matrix providing information, including budget, source of funds, expected start date and use of findings, for the planned evaluations in support of the country programme, for its entire duration. CEPs are kept flexible to adjust to changing circumstances, as for any medium to long-term planning tool; nevertheless, no ex-post mechanism for accountability on the actual implementation of the CEP appeared to be in place.
129. Although no specific guidance exists for CEP development; in general, interviewees at country level appreciated that the process leading to its preparation, sometimes with the support of the Regional Evaluation/M&E Advisers and Specialists, “force” them to think strategically about evaluations as tools that can provide useful evidence in view of, for example, mid-term reviews, efforts to upscale on-going initiatives or the next country programme. Reportedly, CEPs are also useful to schedule the mandatory evaluations requested by donors in the project documents.
130. Interviews suggested that country offices typically struggle between limited human and financial resources, the requirement for evaluations as accountability tools for some donors, the perception that the evaluation process is cumbersome and does not always entail strong added value, and the genuine interest and need for evidence about the results of their work within short time-spans. How this mix of drivers affects the actual evaluation plans, is a management decision of the Country Representative, under the oversight of the Regional Director, advised by the Regional Evaluation/M&E Adviser and Specialist.
131. At the regional level, Regional Evaluation Plans are prepared. The scarcity of human and financial resources emerged as the main constraint to stronger engagement in regional and multi-country evaluations. Opportunism linked to funds availability for certain regional programmes appeared to be the driving factor; under the circumstances, this is probably the best and most pragmatic decision that can be made.
132. With regard to evaluation coverage, the REP identifies four mandatory criteria for planning evaluations to ensure adequate coverage of UNICEF work and performance: programme budgets at or above US\$ 10 million; before scaling up or replicating innovative and pilot initiatives; after 5 years without any evaluation activity; and L3 humanitarian operations. The REP also specifies that evaluations should not be carried out when evidence for making decisions is already available through other sources and assigns to the EO the responsibility to ‘*gauge the coverage and quality of evaluations through a quality assurance system*’. In fact, the role of the EO in the DEF evaluation coverage is one of ex-post monitoring and reporting, largely based on the KPIs and the information that emerges from the GEROS quality assessment mechanism.
133. The provisions for coverage in the REP are therefore quite mild, informed by a strong attention to careful use of resources and to avoiding duplication of efforts. Nevertheless, the only KPI on coverage, ‘Number of countries undertaking evaluation in the last three years’, has shown persistent shortcomings, in particular for large-programme countries, small country-offices and

humanitarian operations. Concerns in this regard were raised several times by the Executive Board, by Senior Management in its responses to the Annual reports on the Evaluation function and by the 2013 Review of the evaluation function.

134. In April 2017, data available in UNICEF corporate systems indicated that, during the period 2013-2016, 29.6% of the countries where UNICEF works did not go through any 'country-focused' evaluation.<sup>39</sup> A majority of countries where evaluations were conducted, 89 or 56% of all those that had 'country-focused' evaluations, went through one evaluation per year at most; and only four countries went through two or more evaluations per year. These figures indicate that although many evaluations are conducted yearly through the DEF, these are highly-dispersed across the whole decentralized network and that a majority of the countries where UNICEF works, either do not go through any evaluation, or through one evaluation only, in a typical country-programme cycle. These data are shown in Box 3 below.

**Box 3. Number of evaluations by country in the period 2013-2016**

Number of evaluations	Number of countries	Percentage of countries within total number
0	47	29,6%
1	38	23,9%
2	22	13,8%
3	20	12,6%
4	9	5,7%
5	11	6,9%
6	4	2,5%
7	4	2,5%
8	1	0,6%
9	2	1,3%
12	1	0,6%
<b>Total</b>	159	100%

Source: GEROS and Country Office Reporting databases, elaborated by the Peer Review

135. At the regional level, during the period 2013-2016, four regional offices issued less than one evaluation per year, and two did not issue any. A notable exception was the CEE/CIS Regional Office which managed six multi-country evaluations in four years. Here, the multi-country evaluation approach successfully enabled addressing strategic themes that were relevant for small groups of countries in the same region that shared similar challenges in collaborating with UNICEF on the same themes. Although the Peer Review heard of countries resisting the approach, there seems to be scope for a broader use of multi-country evaluations across the DEF, supported by relevant advocacy efforts.
136. With regard to the evaluation coverage of humanitarian operations, evaluations of L3 emergencies are a responsibility of the EO, directly or through the Inter-Agency Humanitarian Evaluation mechanism (IAHE), whereas L1 and L2 evaluations are responsibility of country and regional offices respectively. The Annual reports for the evaluation function have repeatedly noted that only a small fraction of UNICEF humanitarian operations have been evaluated so far, also linked to the limited capacity at country level for this type of evaluation. In this respect, and in addition to two on-line training initiatives, the EO launched in 2015 an in-service capacity

<sup>39</sup> As mentioned earlier in the report, work in some of these countries may have been evaluated through a multi-country, regional or global evaluation.

development initiative on humanitarian evaluation, through a coaching and quality assurance process delivered from EO itself. The effort, which will reach in mid-2017 the end of its first round of training, appeared quite effective, despite being resource-intensive, in terms of both high travel costs and EO staff time.

137. Overall, with respect to REP provisions for evaluation coverage, a few observations emerge:
- the REP, or any other document, do not clarify how the established criteria allow an adequate evaluative coverage of the work of UNICEF; for example, no explicit rationale is provided for setting the threshold for project evaluation at US\$ 10 million and what share this represents of the organization's priority programmes or strategies;
  - the information management systems currently in place focus on three of the four established criteria, namely overall evaluation expenditures, number of evaluations per country and evaluations of humanitarian operations;
  - there is no monitoring of the evaluation coverage of innovative and pilot initiatives, nor of any risk category.
138. Thus, on the one hand, there is a gap in defining the benchmarks and targets for an adequate evaluative coverage of UNICEF work; and on the other hand, there are gaps in monitoring the REP criteria for evaluation coverage. Conducting evaluations for the sake of compliance with a KPI in the dashboard may not lead to an efficient nor effective use of resources, as was mentioned in some of the interviews. It is also acknowledged that several country offices do not have the capacity to conduct evaluations. However, it was not clear to the Peer Review which of the evaluation KPIs were followed-up by Senior Management for accountability purposes, as a few country offices also stated that they preferred to conduct no evaluations, rather than have submitted evaluations being poorly rated by Geros.
139. In practice, the list of 'country-focused' evaluations in the period 2013-2016 shows a broad variety in the scope and object of evaluations. Some were pitched at a more complex level of analysis than others, for example were focused on the country programme, or on UNICEF positioning in the country. Conversely, many others were focused on individual projects or global programmes implemented at the national level.
140. In this respect, the ToR for the Peer Review asked to what extent, evaluation topics selected reflect the strategic directions and concerns of the organization. There is no doubt that the strong concern of UNICEF with equity and innovation has been fully integrated in the REP and in the guidance materials issued by EO. At the same time, by definition, evaluation follows the work implemented by the organization. The extent to which UNICEF work at country and regional level reflects the 'strategic directions and concerns' of the organization was, however, an issue beyond the mandate of this Peer Review.
141. At the same time, given the strong decentralized structure of the function and in the absence of an entity in charge of coordination of the evaluative effort, a prevalence of attention on the local over regional and global needs for evidence generation, is to be expected. Moreover, the need for accountability requires extended coverage, country-wise and topic or programme-wise. Nevertheless, the list also conveys a feeling of the 'wheel being reinvented many times', and of missed opportunities for a better organizational learning and a more strategic and rational use of resources. Two examples in this respect were provided. In one case, sectoral reviews of a major humanitarian operation should have fed into an EO-managed broader evaluation, but were not produced in time for this to happen, with the exception of one single review. At the other end of the spectrum, a programme in one country was selected by the EO as case-study for a global evaluation. An EO-selected consultant spent one week in the country, and contributed to the global evaluation that was eventually scored as 'fair' in Geros, which means less than satisfactory. The

Country Office felt they had not contributed to the exercise nor had any ownership for it, and resented being associated with the less than satisfactory rating for the report.

142. There is no doubt that the Evaluation Dashboard and the current KPIs create a useful incentive to conduct evaluations at country level; and that the relevance of evaluations planned very close to where the action takes place is precious. This could nevertheless be maintained in a context where evaluation planning at country level is integrated into regional and through these, global evaluation plans, which would focus on issues having a particular relevance at the corporate level, and for larger number of countries. The information systems currently in place would enable a more coherent planning process for evaluations at the different levels, with the EO in the position to act upon it for coordination and rationalization purposes. Specific suggestions brought forward by interviewees in this sense were as follows:

- enhance communication with the Executive Board, Senior Management and Programme Divisions, to capture their needs in terms of evaluative evidence;
- more inclusive planning of global evaluations with regional and country-level; this would entail the joint selection, sufficiently ahead of the specific global evaluation, of at least some preparatory evaluation work to be conducted at country and regional level; this would enhance at the same time the quality of the evaluative evidence used in global evaluations and the contribution of country- and regional evaluations to learning and accountability at the global level.

143. This shows that there is a perceived need for evaluation plans to ‘talk to each other’ and contribute in a more efficient manner to the evidence requirements at the different levels; the increase in transaction costs during the planning phase would likely be compensated by a more rational use of resources for evaluation implementation and by a better coverage of UNICEF’s work.

144. In the absence of these measures, the current model of evaluation planning at all levels represents a strong limitation to a more strategic and efficient use of the available resources, as well as to broader learning and therefore, to the utility of the function for the organisation. A similar comment was also raised by the Audit Advisory Committee, in its 2016 report, that emphasized the need for the Evaluation Office to have a comprehensive picture of all evaluations planned and completed across UNICEF; and was indirectly also noted by the latest MOPAN report, which stated that *‘the extraction and use of lessons from evaluations is not yet systematic or comprehensive’*.

## **5.4 Evaluation management**

145. Since 2013, the work of the EO in its Plans was first articulated in two ‘streams of work’, then in seven ‘high level results’, and again under two high-level results: Global Evidence and Global Evaluation Governance; and Global Evaluation Management; each with two Intermediate Results, that link the various activities in the Office to the outcome level.

146. All of the EO’s plans have systematically given priority to managing global evaluations. Nevertheless, the actual delivery in recent years fell short of plans, due to delays and over-lapping priorities, for example EO’s engagement in evaluations of humanitarian operations. Further, a few evaluations were cancelled after the evaluability assessments showed their non-suitability for evaluation. Interviews with EO staff also acknowledged that the internal planning process had often under-estimated requirements in terms of staff resources and time to carry out the major global evaluations.

147. In 2015, EO recruited additional evaluation specialists to manage sectoral evaluations, through staff positions and temporary appointments. No additional specific measures were apparently put in place to manage global evaluations in a timely manner, nor for managing expectations of

stakeholders. It is true that global evaluations can be a complex exercise, and that international competitive bidding processes are time-consuming. However, an evaluation report issued 1.5-2 years after initiating the work is likely to suffer from reduced relevance, which may negatively affect its potential utility. Interviews indicated that these long lag-times and delays have contributed to the perception held by some in UNICEF, that the EO would not be able to meet the demand of programme divisions for evaluative evidence.

148. With regards to the management of single corporate evaluations, the EO evaluation specialists affirmed having a certain leeway for deciding the level of their own involvement in the evaluation work proper, including in terms of monitoring and revising contents and quality of evaluation tools prepared by evaluation teams, participation in country visits and direct contribution to the report. Some evaluation specialists have recently started moving towards a management model, whereby they act at the same time as managers and evaluators, in a co-team leadership position.
149. The co-team leadership model has been implemented by other UN evaluation functions;<sup>40</sup> one of its main advantages is that it allows bringing together the technical specialized knowledge of external consultants, and the evaluation experience and competence, as well as the in-depth knowledge of complex organizations, of the staff in the evaluation unit. The main challenge of the model is that its success depends on the collaboration and cooperation between external and internal co-team leaders, which may at times be affected by personality issues.
150. The model could indeed be suitable for the EO, as it would mitigate the refrain heard during the Peer Review interviews, that evaluation consultants have a limited contextual knowledge of the organization. In addition to a different planning of staff-time use, it would require significant adjustments in the tendering documents, with respect to: responsibilities and ownership of the report; time factored into contracts for accommodating a more hands-on EO's role; clear and detailed ToR for both EO manager and team-leaders.
151. Another major issue in the staffing model operational since 2007/08 for the Office in the view of the Peer Review and of some in the EO, was that evaluation specialist positions have been established with funds provided by the programme sections. As of March 2017, approximately one third of the P-positions in the EO were financed through this channel and staff recruited for these positions fully or mostly managed evaluations of programmes and projects under the responsibility of the same Programme section. Interviews pointed to a few risks implicit in this model: potential self-censorship, as by being too critical, the funding for the position could be discontinued; and diminished impartiality, due to long-term relations developed with UNICEF staff who work in the subject area of the evaluations they manage.
152. In addition, staff recruited for these posts have professional background both in evaluation and in the thematic area of the funding programme section. Feedback on this model was mixed. EO staff claimed that, by having technical knowledge of the subject of evaluations, they can: better frame the terms of reference and the evaluation questions; rapidly understand in a critical manner, the information made available by programme sections; develop a network of internal and external contacts that is useful from one evaluation to the next. On the other hand, it was also acknowledged that having a professional stance on the issues at stake, may affect the impartiality of the analysis. Additionally, this prevents internal rotation of evaluation advisers and specialists across EO teams and evaluation subjects, which is likely to prevent cross-fertilization and sharing of experiences over time within the EO itself. Furthermore, because of UNICEF's staff mobility policy which encompasses the EO, staff members might wish to avoid compromising their own chances for career progression into a relevant technical position, which could undermine their impartiality in evaluating the work of the funding Programme section.

---

<sup>40</sup> The evaluation functions in FAO, IFAD and UNDP were broadly in line with the co-leadership model until 2015.

153. Overall, the Peer Review considers that the funding model by programme sections presents real risks to the independence and credibility of evaluations issued by the EO, and for their utility in terms of accountability. With regards to the recruitment of EO evaluation specialists with specific technical competences, arguments in favour did not seem to compensate the potential draw-backs. Although evaluation specialists with a general knowledge of the organization's main subject matters do have a comparative advantage, specialized technical knowledge should be provided by including external senior professionals in the teams, who are constantly engaged in the specific areas to be assessed. Additionally, as is already the case, technical expertise comes through the evaluation advisory or reference panels that are comprised of internationally renowned subject-matter specialists.
154. One of the responsibilities of the EO and of the Regional Offices within the 2013 REP is the development and professionalization of UNICEF evaluation function by promoting internal evaluation capacity development, and by providing guidance and support to the DEF. The EO is the data owner of the UNICEF Evaluation Intranet Web site, which includes a step-by-step guidance on evaluation, from identification and planning to the management response, by cross-referencing to guidance documents of other agencies. Arguably, there is no point in re-inventing work that already exists but this means that the guidance provided, albeit still useful, is very generic and not sufficiently linked to the reality of the DEF. This appears to be particularly challenging for non-specialists, or staff working under time-constraints in isolation. The Programme, Policy and Procedure E-Manual (PPM) does not seem to fulfil this need either, as it basically mirrors the Policy, for which there is no implementation strategy or action plan.
155. Among the various sources of support, a highly-appreciated mechanism for the DEF are the quality assurance systems through LTAs. As of March 2017, five regional offices<sup>41</sup> had contracted, through international competitive bidding processes, consulting firms that coordinate teams of professional consultant evaluators requested to provide real-time feedback on the terms of reference, the inception reports and the draft final reports, for all evaluations in these regions. Thus, country offices engaging in evaluations can still benefit from timely professional evaluation support even when the regional Evaluation/M&E Advisers and Specialists find it increasingly challenging to properly support the evaluations being carried out within their regions, because of time constraints due to multiple responsibilities. Feedback from users was unanimously highly positive, and showed that the initiative would deserve expansion to all regional offices.
156. The interviews of the Peer Review and the 2016 Self-Assessment indicated that UNICEF and UNEG guidance materials and quality assurance systems are appreciated, but that existing training needs do not seem to be adequately met and more guidance is required. For example, the Peer Review noted the absence of standardized templates for evaluation reports which, in addition, is likely to present a strong challenge for the preparation of meta-evaluations and syntheses of evaluations.
157. The 2016 Self-Assessment showed that technical assistance from the Regional Office is requested in more than two thirds of evaluations, for quality assurance of the ToR; and in over half of the evaluations, for the draft report. Linked to these perceived needs, the Self-Assessment also concluded that in UNICEF evaluation function:
- Quality assurance systems are used unevenly through the organization;
  - Quality assurance systems are in place, but they do not always translate into better quality evaluations; and
  - Quality assurance systems and guidance are in place but need enhancement.

---

<sup>41</sup> Namely CEE/CIS, ESARO, LACRO, MENARO and ROSA.

158. Thus, there seems to be a gap in terms of the provision of more precise guidelines on how to conduct evaluations in UNICEF at the different levels. This is needed all the more considering that evaluations are managed by staff who, more often than not, have limited experience in evaluation compared to their other tasks and responsibilities.
159. Furthermore, despite the efforts and achievements in this area of work by both the EO and Regional Offices, as already mentioned earlier in the report, the Peer Review sees a basic flaw in the concept of professionalization of evaluation as conceived in UNICEF. If evaluation is a profession in its own right, it should be assigned to professionally competent staff who have the right skills and experience to discharge their function according to established standards of quality. While there is obviously no harm in sharing information and knowledge on evaluation among UNICEF staff, the cost efficiency of this should be better assessed.
160. Finally, the evidence available to the Peer Review about the profile of the Evaluation Office within UNICEF is rather mixed. Across the various groups of internal stakeholders, including Regional Directors, Division Directors, Senior Management and Executive Board members, different views were expressed. Many stated positive appreciations for EO staff who are regarded as a professional group that works with integrity, and expressed interest and respect for the global evaluations issued by the EO. Several interviewees were, however, more critical of the EO's internal management and considered that global evaluations fell short of their expectations in terms of relevance, quality and timeliness.
161. The strong decentralization of the organization makes the EO and its Director visible and well known in headquarters and at the regional level, but much less so at the country level, in particular for country-level M&E specialists. This is unavoidable, given the limited interactions between the EO and the country offices and the absence of systematic opportunities for a deeper engagement. In addition to the DREAM meetings (see later in the report), the two levels communicate when evaluation teams for global evaluations select countries for in-depth country case-studies, or when an EO staff member participates in reference groups of country-focused evaluations.
162. Stronger visibility and influence would only be possible by providing opportunities for more direct interaction between the EO and the country offices, for example by engaging through the Regional Evaluation Advisers on the Costed Evaluation Plans, including for identifying opportunities for joint evaluation efforts that would be beneficial to all.

## **5.5 Evaluation quality**

163. Ensuring evaluation quality in UNICEF was one of the core elements of the first evaluation policy, with clear responsibilities in this regard assigned to the different organizational levels. The 2013 REP stressed the responsibility of the Evaluation Office for establishing and maintaining a comprehensive quality assurance system to gauge the quality and coverage of UNICEF evaluations. Regional Directors and Country Representatives, in turn, are responsible for the actual Evaluation Quality Assurance process. Attention to evaluation quality is also fully in line with UNEG Standard 5, which focuses on the quality of, and quality assurance mechanism for, evaluations including product and process.
164. One of the main mechanisms in place aimed at enhancing the quality of evaluations, the Long-Term Agreements for quality assurance on the deliverables of on-going evaluations, was discussed in the previous section. The second main tool that is contributing to enhancing the quality of evaluation in UNICEF is GEROS.
165. GEROS is, simultaneously, a systematic quality assessment of all UNICEF evaluation reports and the source of information on the related Key Performance Indicators used for monitoring and reporting on the function. The EO, in its capacity as owner of GEROS, assigns the quality

assessment task to external consulting firms identified through international competitive bidding processes, with triennial contracts. In 2015, GEROS was externally reviewed, which led to significant improvements in the user-friendliness and clarity of the new template used for the assessment, and to some alterations in the assessment scoring system, among other changes.

166. In particular, the new version shows a conceptual shift, from a view where there is clear line between 'good' and 'poor' evaluation reports, to a view of a 'continuum' between the two extremes, as reflected in the change of terminology, from Mostly Satisfactory to Fair. In this respect, a possible option for UNICEF to better capture the continuum across nuances of assessment, in particular when negative, would be the adoption of a six-point scale, similarly to other UN evaluation functions.

167. The GEROS assessment is rigorous and useful, based as it is on appreciative rating, with the double purpose of accountability on the quality of evaluation reports and of capacity development on writing evaluation reports. The mechanism has proved to be quite effective over time in contributing to improving the quality of UNICEF evaluation reports issued by all levels in the function. However, any ex-post assessment of an evaluation report can only focus on the formal aspects and internal coherence of the report itself. There are virtually no ways to identify gaps or misrepresentations in the scoping and terms of reference of an evaluation, unless the reviewer has a good knowledge of the context and of the initiative being evaluated. Whilst, a report can document the process followed; the quality of the interactions, the robustness of the evidence available and eventually considered, and the independence of the process, remain largely unknown to an external reader. This means that a good quality report is not necessarily the result of an equally good evaluation process and GEROS cannot provide feedback on the adequacy of the evaluation coverage; on the integrity and rigour of the process, including in terms of access to relevant information and data; and on the extent of fine-tuning and adjustments of language and contents in the final report, compared to the drafts.

168. With regard to the conduct of evaluations, the steps where quality and credibility are at stake include: the preparation of the terms of reference, the allocated resources, both time-wise and financial, the selection of evaluation consultants, the efforts deployed to enable access by evaluators to complete and comprehensive information, as well as to all stakeholders. Respondents to the 2016 Self-Assessment considered that the function was highly credible, although during the related workshops in the EO and Regional Offices, several factors were identified that would be required for enhancing both quality and credibility, as follows:

- A rigorous evaluation methodology;
- Adequate evaluation budgets, to hire high-quality evaluation teams;
- Sufficient staff capacities (adequate experience/ training) and flexibility to either hire external experts as consultants or conduct parts of the evaluation internally; and
- Availability of practical guidance materials (templates, checklists, etc.), in particular for M&E specialists in the country offices.

169. This list by itself, already proposes a road-map for enhancing the quality of evaluation in UNICEF. Additional direct feedback from interviewees included the following:

- an incumbent Country Representative found it 'odd' that the terms of reference for evaluations at country-level, could be prepared under the sole responsibility of the county office itself, with no independent inputs from outside; admittedly, in many cases the Regional Evaluation/M&E Advisers and Specialists are asked for advice, which is provided either directly or through the established Long-Term Agreements for quality assurance in five regions; as discussed, however, there are gaps in this mechanism;

- financial resources being scarce, consultants may be selected more on the basis of their financial requests than on the strength of their experience and competence in evaluating the initiative at stake; funding limitations may also imply that, in some cases, country offices can only launch tenders at the national level, including in contexts where limited national evaluation capacity is available.

170. With respect to evaluation methodologies, GEROS analysis of hundreds of evaluation reports since its establishment in 2009, indicates that methods tend to be very similar, which is reflected to some extent in the recurrent strengths and weaknesses found in the reports. The 2015 GEROS meta-analysis and other stakeholders also stated that:

- some elements in the reports show persistent weaknesses, including the integration of cross-cutting issues and the quality of recommendations;
- quality of evaluation reports is uneven across regions and strategic plan objective areas, with some systematically better than others; hence, it recommends more internal learning and knowledge exchange on evaluation;
- efficiency is not covered systematically in all evaluations; and overall programme costs were not identified in most evaluation reports analysed by the Peer Review.

171. The 2013 REP makes extensive reference to the importance for UNICEF of integrating into evaluations, the concepts of equity for children, children's rights and gender equality. The templates prepared by the EO for evaluation terms of reference and reports, as well as both versions of the GEROS template, give adequate attention to human rights and gender equality; in contrast, the online step-by-step guidance gives prominence to equity-based evaluations, but does not provide immediate links to existing guidance documents for integrating human rights and gender equality perspectives into evaluation.

172. This gap notwithstanding, GEROS-based meta-analyses show a pattern of progressive improvement in the inclusion of the three principles in evaluation reports, even if these have systematically been among the weaker elements in the reports. The 2015 ratings were all lower than in 2014, due to the change in the assessment methodology (i.e. the quality 'bar' was raised). The 2016 ratings will show whether the previous positive trends will have continued. Key findings, also from the Peer Review itself, were the following.

- Integration of equity as a principle climbed from 9% for 'good' quality ratings in 2010, to 64% in 2014, and 44% in 2015; this is the theme that reached the highest proportion of 'good' quality ratings, in line with UNICEF's efforts and investments around this theme since 2011, including the guidance document on Equity-based Evaluation;
- Good quality ratings for human rights also improved, from 19% in 2010 to 57% in 2014, and 33% in 2015;
- The inclusion of gender in the methodology, framework, findings, conclusions and recommendations also improved from 20% of positive ratings in 2010, to 51% in 2014 and 33% in 2015; the GEROS 2015 meta-analysis also assessed the evaluation reports against the UN SWAP criteria, as adapted by UNEG. The overall result, 'approaching requirements', was consistent with the scoring of '*similar entities that are subject to independent external review or peer review, including UNDP, UN Women and UNESCO*';<sup>42</sup> and
- within the small sample of evaluation reports analysed by the Peer Review, weaknesses were noted in the use human rights language and concepts; also, virtually none mentioned any specific methodological measure to address gender equality concerns.

---

<sup>42</sup> UNICEF GEROS Meta-Analysis 2015.

173. There is thus clear room for improving guidance available to both the EO and the DEF on integrating the three themes in evaluations, with more effort required on human rights and gender equality. In this respect, a particularly interesting initiative in evaluation approaches by the EO, recently started at the time of the Peer Review, focussed on the development of a framework for making evaluations more participatory by systematic engagement with children, throughout the whole process including restitution of evaluation results. Although manuals and guidance on this theme already exist (prepared by UNICEF and other organizations), this exercise aims at developing a complete framework, that will define the necessary political engagement for this to happen, the additional time and financial resources, skills, competences, methods etc. Once completed, the framework will undoubtedly represent an innovative contribution for the organisation and for the international evaluation community.
174. The 2015 Geros meta-analysis reported persistent low ratings over the years for the section on lessons learned and recommendations. Weaknesses identified were mostly due to gaps in the identification of the target groups for action and in the realism of the recommendations, the latter being a typical factor undermining effective follow-up action. Possibly also linked to the low level of realism, the 2016 Self-Assessment raised the request for more action-oriented recommendations, from both directors and M&E specialists. At a more senior level, however, the Peer Review heard requests for recommendations to be more strategic and to provide longer timeframes for their implementation.
175. The system in place should indeed favour enhanced ‘actionability’ of recommendations. UNICEF established that the Management Response should focus on actions that can be reasonably planned and implemented within one year, as anything beyond this time-span cannot be realistically predicted. The approach is sensible, and largely in line with the international guidance in this respect, that stresses that recommendations should be relevant, realistic and feasible. The one-year boundary however, which is associated to a KPI as mentioned above, automatically skews agreement towards those recommendations that are within easier reach, including in terms of tracking as the proof of evidence falls on the implementing unit itself. Recommendations that respond to these needs tend to be mostly operational in nature. Typically, actions and changes at a more strategic and programmatic level take longer and are less tangible in nature and this category of recommendation may benefit from the application of a longer compliance timeframe.
176. The ‘local’ focus of most evaluations in the organization creates a situation where, although decisions, policies and procedures decided at the corporate level may negatively affect programme performance at the local level, few or no recommendations will be formulated for tackling these corporate issues, which are beyond the scope of decision-making of a Country Representative. Only global evaluations or meta-evaluations, if capable of capturing the evidence from the country case-studies or other evaluation reports, can address issues at the corporate level.
177. In this context, a more in-depth analysis of what is considered an adequately actionable recommendation at the different levels of the organization might be warranted to better tackle the existing needs of programme managers and to develop appropriate guidance for recommendation formulation. This should, at the same time, allow space for broader, more strategic recommendations to emerge from the DEF. The development of a recommendation taxonomy could be part of this process, which would also contribute to the analysis of issues and trends emerging from the DEF (referred to later in this report).
178. With regard to the actual conduct of evaluations ‘in the field’, feedback from team leaders of UNICEF evaluations indicated they were in total control of the methodology, as typically requested by the ToR. Overall, the feedback was positive in terms of independence of the process of the evaluations they had led. Nevertheless, a few cases were mentioned of potential threats to independence, which included ‘a push to soften the language’. In one case, programme division

staff in headquarters expected to be part of field visits to communities when conducting a global evaluation; adequate measures proposed by the EO were successfully implemented to avoid any type of undue influence; however, this might not be an isolated, one-off case and the lack of clear guidance on how to handle this type of situation opens the way to potential risks.

179. In all the above situations, the professional ‘clout’ of the team leaders and members was the main safeguard to the integrity of the process. It is not certain that the same respect for the independence of the process would be in place, in the case of evaluations led by less experienced consultants, in an accountability set-up where the evaluator is recruited and responds directly to the head of the office being evaluated, as is often the case for country-focused evaluations. It was also mentioned that in some cases, UNICEF staff did not recognize that an evaluation is a contribution to learning and accountability, separate and distinct from management’s views.
180. The evidence perhaps suggests a limited understanding of evaluation as an independent process to be answered through the Management Response. It also contradicts the standard disclaimer in all UNICEF evaluation reports, which establishes a clear distance between the organization and the contents of the report, while stressing that evaluation fulfils a corporate commitment to transparency and the purpose *‘to stimulate a free exchange of ideas among those interested in the topic and to assure those supporting the work of UNICEF that it rigorously examines its strategies, results, and overall effectiveness.’*
181. Engagement with stakeholders during the evaluation process is a recurrent feature of the 2013 REP, for transparency, credibility and utility purposes. Stakeholders are identified internally and externally to UNICEF, across the different levels of the organization, within national governments and among other partners. The on-line step-by step guidance dedicates a section to stakeholders’ involvement, with no specific reference to, nor template for, providing feedback and comments in the report. Hence, it is inferred that each office would handle the process according to its own circumstances.
182. This emerged as a weakness in the feedback from evaluation team leaders, some of whom were quite critical of the modality for canvassing comments from stakeholders. The main points raised were the following: drafts were circulated for comments to all stakeholders at the same time, whereas in other agencies, it is usually a phased approach across different levels of engagement and oversight responsibility; a lack of structure in providing comments; the absence of a prioritisation within comments, whereby minutiae on wording was more prevalent than strategic and substantive comments.
183. In addition, there was no clarity in the number of iterations through which a report was supposed to go, which resulted in large amounts of time spent on minor issues. Although it was acknowledged that comments overall contributed to improving the quality of evaluation reports, the lack of clarity on boundaries between evaluators and management mentioned above, and the dilution of the strength and core messages of an evaluation report through successive layers of comments, did not necessarily contribute to the credibility of the reports.
184. Related to this, and confirmed by interviews, the Peer Review noted a very strong corporate culture to pursue outcomes for the benefit of children and to refrain from disclosure of less than optimal results. However, the unchallengeable commitment to the mandate must not be used as a justification that usurps attention to accountability requirements or learning opportunities, even those that arise from negative evaluation evidence.
185. The REP identifies as good practice, establishing a stakeholders’ reference group early in the evaluation process and both Reference Groups and Evaluation Management Groups appear to be established for most, if not all evaluations across the organization. Membership of the groups varies and may include country representatives in the case of a multi-country evaluation; or the regional

M&E specialist in the case of country-level evaluations. In one case, it was appreciated that it also comprised one internationally-renown expert in the technical area being evaluated. Still, at times composition may not follow the basic rules of segregation of duties, e.g. mention was made of a Reference Group incorporating the evaluation manager. Or in small country offices, the Evaluation Management Group might comprise the section chief of the programme being evaluated.

186. Overall, the use of reference and advisory groups seems to be well embedded in the practice of evaluation in UNICEF. However, more guidance would be welcomed, clarifying the roles of each of these groups and the best practices for their composition.
187. Comments were also raised, related to the competences of the UNICEF evaluation function on experimental and quasi-experimental evaluation methods, in particular throughout the DEF; and on the methodologies used by the evaluation function for this category of evaluations. In this regard, a reference to the broader picture on impact evaluation across the UN evaluation functions may be useful. Competences in these methodologies tend to be scarcer across UNEG members, compared to competences in qualitative and simpler quantitative methods. As determined during the data-gathering phase for a UNEG guidance document issued in 2013 on impact evaluation in the UN system, this scarcity was also due to the limited usefulness and applicability of experimental and quasi-experimental methods for impact evaluation, to the mainstream work of the UN.
188. The Agenda 2030 may appear as a game-changer in this respect. Although the work of the UN will increasingly move more upstream, towards the policy and institutional development level, governments will increasingly need more statistically robust information on the impacts of their own policies and programmes, with responsibility for this type of analysis likely to fall on national statistics systems. Arguably, support to the latter would appear to come more appropriately and effectively, from the technical and programmatic arms of the UN, than from evaluation functions that should primarily serve purposes of internal learning and accountability. To some extent, within UNICEF this scenario is already in place, with the Office of Research taking the lead on impact assessments through a Long-Term Agreement for this type of analysis. This could indeed be formalized at the corporate level, as a clear division of roles and responsibilities between Evaluation and Research also in terms of support to country offices that want to launch experimental or quasi experimental impact assessments. This would also help to clarify the specificity of evaluation's contribution to the evidence-generation function.
189. The evidence above shows a mixed picture on the credibility and quality of the evaluation process at country level. In the case of global, regional and multi-country evaluations, the majority of the steps in the process are managed in a context of stronger independence, more resources, better competences and experience in evaluation. However, country-level evaluations form the large majority in the organization, which calls for some urgent re-thinking of the model, if UNICEF wants to be appreciated also for the quality and credibility of its entire evaluation function.

## **5.6 Evaluation follow-up and use**

190. The importance of evaluation follow-up and use has been a recurrent topic in the deliberations of the Executive Board on evaluation, supported by Senior Management. The 2008 UNICEF Evaluation Policy already included among the Guiding Principles for the Evaluation Function that Management Responses (MRs) to evaluations should be prepared and made publicly available and the Executive Board requested several times more compliance with this requirement. This led over the years to a marked improvement in the rate of MRs submitted, reaching 97% of the evaluations in 2015.
191. This strong attention by UNICEF governance and Senior Management to Management Responses and their follow-up, appears somewhat in contradiction with the decision to discontinue

public access to MRs, as was the case with the 2008 Evaluation Policy, all the more so having the added tool of GEROS assessment on the reliability of an evaluation report. The advantage of such a mechanism is that, in case of poor quality evaluations or recommendations,<sup>43</sup> Management can publicly protect its own image through the MR, without interfering with the independence of the evaluation process.<sup>44</sup>

192. In early 2016, the Evaluation Office issued the document 'Guidance for Management response to evaluations', which clarified the process and the different responsibilities in the Management Response process. The MR preparation is responsibility of the head of the office, whereas programme heads are responsible for developing and implementing the agreed recommendations through identified actions. M&E staff support to the MR process in more than two thirds of the evaluations. The guidance allows two months for the preparation of the MR after the finalization of the evaluation report, during which extensive consultation can take place with all concerned stakeholders on the actions to be planned to respond to the recommendations.
193. The team responsible for the preparation of the MR should also receive in a timely manner, the GEROS assessment on the quality of the report, that "*can help the stakeholders have confidence to accept or reject major elements of the evaluation*". This should help in focusing action and follow-up on the 'good quality' recommendations, rather than accepting all recommendations as was found by the Review of the Evaluation Management Response mechanism completed in 2016. However, given the limitations inherent in the GEROS assessment, a 'good quality' recommendation may not be the most relevant or appropriate. In practice, the GEROS assessment may be useful to users to decide on whether a recommendation should be accepted, as long as this step does not delay the MR reparation process, which would lead to a loss of the momentum generated by the evaluation process.
194. The 2013 REP states, among the performance standard for the function, that evaluations must be used to the maximum possible extent and that "*Country Representatives, Regional Directors and, as appropriate, Divisional Directors will ensure that recommendations are fully considered with concerned partners, that accepted recommendations are acted on, and that annual reports include a statement on the status of evaluation follow-up*". This commitment is operationalized through the KPIs that cover, respectively, the number of evaluations complemented by management responses, and the implementation rate of the management responses.
195. In compliance with this provision, UNICEF established the Evaluation Management Response tracking system (EMR), a corporate system where all recommendations for each evaluation are recorded, and regular updates are made of the state of progress in implementing them. Follow-up on the actual implementation of recommendations is under full control of each responsible office, which uploads information into the EMR system on the rate of agreed or rejected recommendations and on progress in implementation. No mechanisms are in place for external validation.
196. Moreover, the EMR does not include a taxonomy for different categories of recommendations, which would enable extracting consolidated data across all records, about the degree of agreement and extent of progress in implementation of, for example, recommendations addressing policy or strategy development. Admittedly, even should such a taxonomy be in place, more detailed analysis would be necessary to understand exactly what has been achieved through the follow-up actions. Thus, it would be the responsibility of the evaluation managers to highlight recommendations or actions that achieve certain types of impacts.

---

<sup>43</sup> This refers to those cases where the recommendations are not relevant, not realistic, or do not follow logically from the evidence gathered.

<sup>44</sup> In other UN organizations, this mechanism is taken very seriously, to the point that global evaluation reports are not made publicly available until the Management Response is also available.

197. At the same time, the absence of a link between Geros and EMR does not allow lessons to be drawn on whether any relationship exists between the formal quality of a recommendation and the extent of its implementation. Arguably, the Management Response Review carried out this analysis did not find any strong proof of evidence in this respect.
198. The Annual reports on the Evaluation function have addressed the use of evaluations, and EO launched surveys that tackled this issue among others. The 2013 Review of the evaluation function indicated that “*Evaluation findings are incorporated into evidence-based policy advocacy (indicated by 53 per cent of respondents as a “high use”) and programmatic decision-making (42 percent)*”. More recent analyses of this type were not conducted and within the current systems, therefore, the only way to gain insight into the impact of evaluations is through anecdotal evidence compiled through ad-hoc efforts, as also acknowledged in the Annual report for 2014 on the evaluation function in UNICEF.
199. Feedback to the Peer Review from interviewees in headquarters suggested that most global evaluation reports are taken seriously by Senior Management and are regarded as a source of credible evidence. It was also stated that Country Representatives have used evaluation reports for advocacy purposes with donors and governments and information was shared with the Peer Review, about uptake of evaluation findings in national policies and programmes.
200. Other evidence includes an impact evaluation that adopted a quasi-experimental design, more expensive but considered more robust to assess changes in the behaviour of mothers that participated for several months in a Communication for Development (C4D) project to improve maternal and child health and which led to upscaling the programme to the national level. Mention was also made of country strategy and programme evaluations that contributed to the preparation of new Country Programme Documents and to better positioning the organization in the national institutional context.
201. However, no systematic mechanism is in place, nor analysis conducted to capture the long-term impact of evaluations on influencing and learning, within and outside the organization. The initiative that comes closer to effectively sharing lessons and experiences from evaluation with other staff in UNICEF are the annual Data, Research, Evaluation, Analysis and Monitoring (DREAM) meetings, organized by various offices in UNICEF, with the active contribution of the EO. These are considered very useful to discuss the links between all the pillars contributing to the evidence function, but are not the right venue for discussing the impacts generated by evaluations.
202. UNICEF is not alone in facing this challenge, shared by all UN evaluation functions. Going beyond the anecdotal evidence requires time and resources, both from the monitoring staff and Performance Management Teams, who need to establish system for tracking progress in implementing agreed recommendations over longer periods, and from the evaluation users, who have to provide the ‘proof of the evidence’. Moreover, there are many variables along the impact pathway of an evaluation, which may undermine the usefulness of the evaluation and its recommendations due to factors completely out of control of the evaluation manager and the head of the office, let alone the evaluator. At the same time, reporting on implementation progress of agreed recommendations may also be subject to some bias; and this type of reporting, to be fully credible, requires some form of independent validation.
203. Communicating evaluation results is another challenge common to most evaluation functions in the UN system, and no easy recipes exist. One of the key findings of the Self-Assessment was that ‘*Use by affected population and the general public is low and few evaluations use innovative formats to communicate results*’. Typically, in addition to the reporting to the Executive Board, UNICEF organizes the stakeholders’ workshop to discuss the final report and its recommendations, in view of the preparation of the Management Response, and the uploading of the report in the Web

site. The same would apply to joint evaluations, with adjustments made according to the procedures and preferences of the partners.

204. The EO has issued four syntheses of evaluation reports in the period 2013-2016, to make results from larger bodies of evaluative evidence available in a more succinct and digestible format. Identifying additional options for improving the dissemination of findings and lessons is the task of the System Strengthening Unit, which recently recruited a Knowledge Management specialist, at P3 level, to fill the existing position that had remained vacant. This should help to improve the visibility of evaluation in the internal corporate channels for knowledge exchange, and enhance the production of evaluation knowledge outputs.
205. Within the DEF, it is reasonable to assume that staffing capacity will be the main constraint to proper work on dissemination of evaluation findings and lessons. At the country level, given the various responsibilities of M&E specialists, country offices that include Communication specialists among their staff who can contribute to this effort, are likely to be in a better position to disseminate lessons learned from evaluations. At the regional level, although a few good examples exist, arguably more could be done without significant additional resources, by disseminating evaluation findings and lessons emerging from country-level evaluations, among regional specialists in any given programme and among countries in the same region, which may be facing similar challenges. Some stakeholders also suggested the potential usefulness of two-page briefers on key evaluation findings.
206. With respect to the corporate capacity for evaluation absorption, the relatively high number of evaluations issued by UNICEF on an annual basis, around 100 as an order of magnitude, must be compared with the strongly decentralized structure of the function. As discussed earlier, the majority of countries go through less than one evaluation per year, which should not be a challenge. At the global level, the EO has been issuing on average six evaluation products per year over the last four years, which are not few but should not be a threat to the absorption capacity of an organization like UNICEF, considering that Senior Management and the Executive Board had endorsed the Global Evaluation Plan which foresaw nine evaluation products/year. Moreover, as confirmed by some interviewees, the fact that some programme divisions have established their own evaluation units, indicates that there is an un-met demand and need for evaluation in the organization.
207. The extent to which the results of UNICEF's evaluations contribute to the broader process of knowledge management, is more complicated to assess. The organization gives strong attention to the knowledge agenda and has a number of Web pages in its Intranet specifically enabling knowledge exchange.<sup>45</sup> A rapid view of the Web page 'New & Noteworthy' showed very few references to evaluation, which seemed to be a missed opportunity. In line with this, most interviewees of the Peer Review mentioned an existing gap in managing the knowledge generated by evaluations.
208. The 2013 review of the evaluation function had found that "*Knowledge generation was the evaluation purpose referenced the least often; those who did mention it mostly discussed knowledge generation as a by-product of evaluations rather than as their primary purpose.*" It also found that the main use of evaluations was decision-making and evidence-based policy advocacy, as confirmed by the 2016 Self-Assessment. In addition to challenging the corporate discourse and policy that define evaluation as one of the pillars of the evidence-generation function, these findings would suggest little interest and possibly limited intention among directors, representatives and chiefs of units, to 'manage the knowledge' generated by the evaluations they are accountable for.
209. The contradiction emerging from the evidence above might stem from the following causes:

---

<sup>45</sup> Knowledge exchange is a necessary condition, though not sufficient, for knowledge management which should entail a more complex process of evidence analysis and synthesis, followed by preparation of communication products and their diffusion.

- The decentralized planning of evaluations leads to evaluations being of high relevance at the local, country office level; as Evaluation Management Teams comprise programme and office management, the learning process runs in parallel with the evaluation itself and once the report and management response are finalized, there is likely limited need and incentive for additional efforts in sharing knowledge;
- Knowledge exchange and management across country offices, or across regions, requires an entity at the regional and global levels, that not only extracts information from evaluation reports, but also steers and coordinates the type of information and evidence that would be interesting and useful to share at a higher level; so far, UNICEF does not have such an entity nor role assigned to any level in the organization.

210. To some extent, the challenge represented by Knowledge Management in the evaluation function must be recurrent in UNICEF, given the strong decentralized structure and culture of the organization that will affect all efforts to share any type of knowledge across regions, and possibly also within each region. In this context, the pursuit of any improvement towards enhancing the utility of the evaluation function above the country level, could rely on lessons learned on knowledge management within the organization, if any exist, that could be useful for the evaluation function. An alternative option would be to revise to some extent the mandate of the Evaluation Office, giving it a role in evaluation planning, conduct and knowledge management across regions and countries.

211. Partly linked to the use of evidence generated by evaluations, a recurrent issue raised with the Peer Review by interviewees, concerned the reported gap in evidence available to managers on the results of their work at the outcome and impact level. The challenge they face is how to get the type of information that evaluations can generate not just towards the end of a programme, but continuously throughout implementation so as to be able to adjust and steer the programme accordingly.

212. This would be, in the understanding of the Peer Review, the role of outcome and impact monitoring, to be conducted as a management tool (internal and light). This might require additional resources for programme monitoring and should be based on the key indicators identified in the Country Programme Document, for this level of assessment. The new tool developed by the Field Resource Group seems to be a good step in this direction. If this type of monitoring were to be established and managed by the monitoring specialists in the country offices, results would feed into evaluations proper, later in the implementation cycle, which could focus on a higher level of analysis with basis on robust monitoring data about results, not only delivery.

213. Last but not least, in recent years, UNICEF evaluation reports have also been used by other partners, namely the MOPAN and an OECD/DAC Evalnet-led assessment of UNICEF, as evidence base for their own analysis of the performance of the organization. The OECD/DAC-led exercise started with a quality assessment of a sample of UNICEF evaluation reports, to validate the accuracy and reliability of the GEROS quality assessment. The two systems coincided to a very large extent, 95%, which allowed a systematic use of the evaluation reports as source of evidence for the broader assessment.

## **5.7 External influence, partnerships and positioning of UNICEF evaluation function**

214. UNICEF values both Capacity Development (CD) and Partnerships, which are two of the seven implementing strategies of the SP and are important elements of the external projection of the organization. The 2013 REP fully embraces the corporate commitment, integrates both strategies into the organizational mandate for evaluation and assigns responsibility to the Evaluation Office for the area 'Partnership for Evaluation', which entails "*promoting national*

*ownership and leadership of evaluation activities; supporting global partnerships and networks; and working through these to support innovation and evaluation-capacity development; engaging with UNEG to harmonize evaluation norms and practices and to build United Nations evaluation capacity; and contributing to United Nations system-wide evaluations”.*

215. In addition, the REP discusses National Evaluation Capacity Development among the standards for evaluation, in consideration of the significant role UNICEF had played in the past in this respect and of its continuing commitment to this goal. The REP also acknowledges that the way forward on NECD can only be in partnership with the UN system, and identifies five possible categories of work for UNICEF offices to engage with. These include both programme-type interventions, as well as initiatives that can be easily embedded in evaluation management and approaches.<sup>46</sup> Strategies were to be developed to tailor the two streams of work to the different circumstances.
216. However, the REP does not clarify who, how and with what resources, the programme-type interventions in support of National/Evaluation Capacity Development (N/ECD), should be conducted. In the absence of a strategic action plan for the REP, and for NECD, two key questions remain un-answered: How can UNICEF M&E specialists at country level who, in general, do not have a strong background in evaluation, develop capacities at the national level in evaluation? And who will evaluate NECD work, if EO staff are heavily engaged in conducting it?
217. At the corporate governance level, although the Executive Board over time did encourage and request that due attention be given to NECD, it did not appear to consider NECD as a top priority for UNICEF; rather an area of work where efforts should be coordinated among UN agencies to avoid duplications and overlaps, with Resident Coordinators playing a leading role.
218. With respect to N/ECD, the EO has been very active over the years in producing several key guidance books and material, often in close partnerships with other agencies, either directly focused on how to conduct N/ECD or contributing through information and experience sharing, to N/ECD. During the time-timeframe covered by the Peer Review, this type of work slowed down to some extent.
219. Within the DEF, the Peer Review heard anecdotal evidence of challenges and successes in NECD, in developing partnerships with governments and in country-led evaluations. These included cases where national evaluation units were being established in government organizations, as well as national contexts where an independent evaluation could not be conducted or made publicly available. In early 2017, the EO launched a mapping of UNICEF NECD activities at all levels. This is a praiseworthy initiative, which will represent a baseline for this area of work at the corporate level, on the basis of which well informed decisions in terms of future engagement can be made.
220. Finally, feedback from the interviews and the Self-Assessment on UNICEF’s role in NECD was mixed. Most stakeholders recognized the comparative advantage of UNICEF at country level, thanks to its network of M&E specialists and to its engagement at the global level. However, most agreed that for UNICEF to make progress in this direction, a common strategy was required. In this respect, different opinions were held, ranging from the belief that support to national and regional evaluation associations is cost-effective, to the view that NECD is best achieved by doing evaluations together with national partners. In either case, a fundamental pillar of the strategy should be the separation of the evaluation of the results of NECD from those actually responsible for its implementation.

---

<sup>46</sup> Programme-type interventions include ‘Establishment and strengthening of national evaluation systems’, ‘Strengthening of sectoral monitoring and evaluation or management information systems’ and ‘Training in monitoring and evaluation concepts and use of evidence’. CD initiatives embedded in evaluations are: ‘Inclusion of NECD objectives within major programme evaluations’ and ‘Advocacy for and investment in country-led evaluations as well as use of evidence from such evaluations’.

221. In the view of the Peer Review, there is no doubt that UNICEF experience in partnership development and on NECD are precious assets that should be brought to bear into the on-going debate within the UN evaluation system, about how best to address the Agenda 2030 challenges on country-led evaluations and NECD. In particular, UNICEF can provide to the UNEG debate on NECD, valuable lessons on which NECD activities can be carried out through a three-tier evaluation function, how effective and efficient it can be to give direct responsibility for NECD to country offices; and under which circumstances partnerships with governments for independent and rigorous evaluations can be a fruitful effort.
222. With regard to partnerships, the EO directors have systematically been actively engaged in networking and participating in external events and fora. Accordingly, the Annual reports on the Evaluation Function, and the EO Annual Reports, have given high visibility to the EO's contribution to UNEG. In recent years, these activities included hosting the annual UNEG Evaluation Practice Exchange event and leadership of thematic working groups. One of the EO's Senior Evaluation Specialists is formally responsible for coordinating this area of work, while EO staff, both senior and more junior, have been members and co-chairs of UNEG working groups. Available feedback from staff, did not suggest that participation in these networks was excessive compared to their other duties. Furthermore, some of the Regional Evaluation/M&E Advisers and Specialist have been contributing to regional evaluation networks.
223. Additionally, the EO played a major role in the establishment of EvalPartners, a 'global partnership to strengthen national evaluation capacities', which is contributing to supporting NECD in many countries, stimulating debate on and developing innovative approaches for evaluation. The Director of EO at the time of the Peer Review, had also taken on the role of UNEG Vice-Chair and held the role of Vice-Chair of EvalPartners; in this position, he contributed to the launch of the EvalSDGs Network. EvalPartners highly appreciated and valued the support received, which strongly contributed to giving visibility and credibility to the network. In the view of the Peer Review, however, the commitment to these external networks may have come at a cost for the work of EO itself, in particular with respect to timeliness of global evaluations delivery and to fulfilling the internal normative guidance role.
224. The EO is also a member of the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP) and of the Inter-Agency Humanitarian Evaluation management group, both deeply engaged in coordinating, managing, and supporting capacity development in the case of ALNAP, in the field of humanitarian operations evaluation.
225. Within the evaluation international community, the EO is well known and appreciated for having spearheaded the concept of equity-focused evaluation, in line with UNICEF's focus on equity, and for developing relevant guidance. The EO has also been at the forefront of EvalPartners, the innovative and unique partnership among evaluation stakeholders from many different venues to strengthen evaluation capacity world-wide.
226. With regard to partnerships in conducting evaluations, data in the GEROS database indicate that between 2013 and 2016, a total of 67 joint evaluations were completed, representing 18% of the total. Of these, three were joint evaluations within the framework of the Inter-Agency Humanitarian Evaluations (IAHE) mechanism. With respect to the latter, UNICEF country offices also conducted evaluations of the same humanitarian operations as those covered by the IAHEs, with a different focus as typically happens for other evaluation functions contributing to the IAHE mechanism.
227. The highest share of Joint Evaluations was conducted in countries in the ESA and in the EAP regions, at 25% of the total number of evaluations in each region. In other regions, the percentage values varied between 12% and 20%. The data show a tendency at having more than one Joint Evaluation in the same country, which is understandable as joint evaluations are likely to reflect a

habit of UN agencies working together in that country. However, no clear pattern could be discerned, not even with respect to the first group of countries piloting the Delivering-as-One approach. Thus, the only conclusion that can be drawn is that UNICEF participates in a reasonable number of joint evaluations. However, the Peer review could not assess whether the current engagement in joint evaluations reflects the extent of joint work by UNICEF with other agencies, nor whether proposals to carry out more joint evaluations were turned down, at any organizational level.

# 6 ASSESSMENT OF THE EVALUATION FUNCTION IN UNICEF AGAINST THE THREE CORE CRITERIA: INDEPENDENCE, CREDIBILITY AND UTILITY

## 6.1 Independence

228. The ToR for the Peer Review articulate the principle of Independence in two key aspects, the behavioural and the organizational independence. The former refers to ‘the ability to evaluate impartially without undue influence by any party’; freedom from conflict of interest and undue pressure, and from concerns about possible future effects on one’s career, are key elements to achieve behavioural independence. Organizational independence, in turn, requires that *‘the central evaluation function is positioned independently from management functions, carries the responsibility of setting the evaluation agenda and is provided with adequate resources to conduct its work’*. The ToR also state that *‘the activities of the Evaluation Office can be expected to have greater degree of independence than evaluation activities at decentralized levels’*.
229. With respect to the behavioural independence of the evaluation function in UNICEF, the Peer Review identified incipient mechanisms that are conducive to strengthening it, namely: the establishment of multi-country evaluation specialists positions, with multiple reporting lines that greatly reduce the possibility of undue pressure and avoid evaluation being one among several roles and responsibilities; and the conduct of multi-country evaluations, that meet country offices’ needs in terms of evaluative evidence through evaluations managed above the country level. With respect to the direct reporting of M&E specialists and Evaluation Advisors to Regional Directors and Country Representatives, this should be the minimum requirement for independence of the function, though the Peer Review acknowledges that within UNICEF, this represents a significant improvement and appreciates that it can serve as a paragon for all other offices where this arrangement is still not applied.
230. Another measure that contributes to strengthening the behavioural independence in evaluations across all levels is the systematic recruitment use of external evaluation consultants to conduct evaluations. This is an overall positive approach and safeguard, although it may not always represent a guarantee of independence and impartiality, in particular in those circumstances where consultants operate in a vacuum of direct contacts and references beyond the commissioning and the programme staff, who frequently work together.
231. At the same time, the accountability framework and the ‘shared’ characteristic of the evaluation function in UNICEF lower the boundaries of evaluation as an independent function as part of the corporate oversight mechanism. The Peer Review is fully aware that this overarching characteristic of the function is deeply embedded in the culture of the organization and acknowledges that it enables a reasonable commitment to the function and use of evaluation. Nevertheless, this set-up is not in line with the international evaluation standards, as embedded in the UNEG Norms and Standards; and has a bearing on the overall independence, and credibility, of the function.
232. Evidence available and discussed shows a number of additional limitations to behavioural independence, mostly but not exclusively in the DEF, as follows:

- a. the reporting lines of Evaluation Advisors and Specialists in Regional and Country Offices, and of evaluation consultants recruited to conduct the evaluations, to senior staff who have responsibilities in the planning and conduct of the initiatives being evaluated;
- b. the financing in the EO and in the DEF of M&E positions from Programme Resources, with staff recruited for these positions who commission, manage, and have an active role in, evaluations of the programmes themselves;
- c. the multiple roles and tasks assigned to Evaluation/M&E Advisors and Specialists in Regional and Country Offices, that blur the boundaries between evaluation and monitoring, planning and research;
- d. the absence at the policy level of practical safeguards for human resources in evaluation, including on staff rotation policy.

233. In terms of organizational independence, the strengths of UNICEF comprise the Evaluation Office, which is independent, well-staffed and resourced and firmly established in the central structure of the organization; the strong commitment of Senior Management to the allocation of 1% of the overall programme expenditure to the evaluation function; and the systematic requests of the Executive Board for an independent evaluation function and its endorsement of significant additional Regular Resources allocated to the EO for evaluation work.

234. The REP however only guarantees the independence of the EO, which suggests more limited corporate attention to the overall independence of the function, despite the fact that the DEF produces the large majority of evaluations. Other identified gaps, real or perceived, in the organizational independence include:

- e. the “day-to-day” reporting line mentioned in the REP of the Director, EO, to a Deputy Executive Director;
- f. the lack of safeguards against potential conflicts of interest, and of external professional evaluation inputs, in the selection process for the Director of the Evaluation Office.
- g. the absence of clear criteria for the selection of evaluation reports to be presented to the Executive Board, and the lack of systematic engagement with the Board to canvass views on evaluative demand;
- h. the accountability framework for key responsibilities and accountabilities in the evaluation function, assigned to directors at the different levels, who combine programme planning and management, and evaluation oversight duties;
- i. the reduced role of the EO in the governance of the DEF, with respect to the situation under 2009 Directive on Evaluation;
- j. the absence of public access to the Management Responses to evaluations, which inhibits at the same time the transparency of the evaluation process and the possibility for the organization to publicly defend its reputation from inaccurate or impartial evaluations.

235. Overall, the Peer Review’s assessment of both aspects of independence is short of being fully satisfactory, both for the EO and for the DEF. The strong corporate culture of ‘doing for the benefit of children’ was very apparent to the Peer Review. This admirable culture and focus also provides a strong justification for robust and independent evaluations to ensure formal accountability needs are met.

## 6.2 Credibility

236. The definition of Credibility in the ToR grounds this principle in '*independence, impartiality and a rigorous methodology*', in the ethical conduct of evaluations, in the professionalism of evaluators, in the participation of stakeholders and in the shared ownership of the evaluations process.
237. The evaluation function is overall regarded as credible within UNICEF, by both users and staff directly involved in managing evaluations. This is partly a reflection of the commitment and energy dedicated to evaluations by staff directly and indirectly engaging in evaluations. Another positive feature of the UNICEF evaluation function that contributes to its credibility is the good level of commitment in the policy to, and the actual engagement with, stakeholders in the evaluation process. Also, over time there has been an improvement in the quality of evaluation reports, as reported through Geros, which has strengthened the credibility of the function. With respect to ethical considerations in conducting evaluations, these appear to be reasonably well integrated in the evaluation process, albeit with some room for improvement, also as a reflection of the strong emphasis given to equity by the EO in the guidance documents issued.
238. On the other hand, the weaknesses identified on independence and impartiality also affect the overall credibility of the UNICEF evaluation function, as discussed throughout the report where independence and credibility are frequently associated. For example, the attribution of the full responsibility for approving IMEPs and CEPs and for managing an evaluation process to the head of the regional or country office, are gaps in the independence of the evaluation process that automatically weaken the credibility of the plan and indirectly, of the end-products. This also means that most measures aimed at improving independence and impartiality of the evaluation process, would also enhance the credibility of its results.
239. With regard to the professionalism of the function, the evidence available showed that most staff responsible for the DEF do not have sufficient experience to make well-informed decisions on technical evaluation approaches and methods and have no time to keep up-to-date on these. The findings through the Geros assessments also indicate that a restricted range of methods and tools are used. Credibility in evaluation is also a function of the quality of the process, and the 2016 Self-Assessment identified a number of gaps in the latter.
240. At the level of the EO, the use of a limited set of methods and tools can also represent a risk for credibility, although there is also evidence of efforts towards an improvement in this regard, namely for a more inclusive and participatory child-focused evaluation framework, which is appreciated and deserves encouragement. Another positive initiative in this sense, if the issue of funding resources for the positions of evaluation specialist in the Office can be resolved, is the proposed model of co-leadership for corporate-level evaluations. This would allow deepening the knowledge on the organization brought to bear in the evaluation evidence and analysis and would strengthen its credibility, in particular in the eyes of internal stakeholders.
241. One important element of credibility that was not included in the ToR, is the evaluative coverage of the work of the organization. The existing KPI on number of country offices that have conducted an evaluation over time, even when crossed with other KPIs, does not provide any insight on whether the completed evaluations have tackled the most important initiatives, or those at biggest risk, or those that might have more potential for lessons learning or up-scaling. The absence of an analysis of what would be a reasonable coverage of the work of UNICEF affects the credibility of the function as a whole, and at each level where it is managed.
242. Overall, the Peer Review's assessment of credibility of the UNICEF evaluation function is short of being satisfactory, both for the EO and for the DEF. Specific actions on the quality of the evaluation process could nevertheless be implemented with a direct effect on credibility, also independently from actions aimed at strengthening independence. At the same time, any action

aimed at enhancing the independence of the evaluation function will also positively affect its credibility.

### **6.3 Utility**

243. Finally, the ToR of the Peer Review define Utility by the '*clear intention to use*' the evaluation results. Key factors in Utility are timeliness of the process, credibility and relevance of the products, as well as commitment and interest of recipients to accept the findings, analysis, conclusions and recommendations. The broader contribution of evaluation to knowledge is also mentioned.
244. The REP is very clear in stressing the importance of the use of evaluations, and underlines that both partnerships and involvement of stakeholders in the evaluation process from its inception, are strong factors in fostering ownership for the evaluations and should contribute to enhance their usefulness both in terms of the relevance of findings and recommendations and also the learning that result from a well-run process.
245. The REP also establishes a strong link between evaluations and the corporate evidence-generation function. Indeed, the stress on the shared nature of evaluation and the association of evaluation with research, largely inform the corporate perception of evaluation as a useful management tool for decision-making and learning. This blurring of purposes however, compromises the utility of evaluations that stems from the independence and impartiality, and therefore the credibility, of the evaluation process. In the absence of the latter distinctive elements, the usefulness of a process that produces outputs similar to research, studies and reviews, but is more cumbersome and may not bring much additional added value, could be easily challenged.
246. In the current practice of the function at the corporate level, there is evidence that some evaluations that were considered highly credible, have been intensively used and mined for advice. At the same time, a few significant challenges to the potential utility of evaluations were identified, including the timeliness in the delivery of the reports, the quality of the process and product, and to some extent, the topics being evaluated. Thus, there seems to be demand and interest for corporate level evaluations that better meet the needs of stakeholders in terms of topics, shorter time-spans for the process, and possibly deeper levels of analysis and use-oriented recommendations. If these needs are addressed, then evaluations will gain in relevance and utility.
247. With regards to the utility of evaluations issued through the DEF, there seems to be two possible opposing situations. In the first, the embedded nature of the evaluation process in the country office structure ensures its high relevance and ownership; this should also naturally lead to a good level of immediate and direct learning from the evaluation process, and to a high utility at local level. In the second, evaluations are conducted solely to meet donors' obligations or to fulfil Dashboard and KPI requirements and become a 'ticking the box' exercise; under such circumstances, their utility might be rather low. Obviously, between these two extremes, a whole gamut of possible mixed scenarios is likely to exist, but no information about which side of the spectrum prevails was available.
248. Evidence available suggests that the stronger utility of evaluations in UNICEF is as advocacy tool, for example with governments and partners; and for decision-making, for example when preparing new Country Programme Documents or expanding pilot programmes. Lessons learning did not emerge as a priority use for evaluations, although it can be easily argued that lessons learning is a first step for both advocacy and decision-making. Nevertheless, the lesser attention to learning affects the potential contribution of the function to corporate level knowledge management, which is, in any case, highly complex due to the strongly decentralized nature of the function. This means that evaluations could contribute to corporate knowledge management if links were created

between the function and the existing relevant internal mechanisms; currently, there is a gap in this respect.

249. In the light of the above, and in the current policy framework, the Peer Review assesses the utility of evaluations in UNICEF to be close to satisfactory, though not quite there yet. Efforts for improvement in scope, timeliness and quality are required at the level of the corporate evaluations.

250. Above all, however, the Peer Review argues that a better definition is required, of what is the most important role of evaluation within UNICEF, and what needs and gaps it must respond to. This should be part of a process leading to the development, in a consultative manner throughout the organization, of an agreed Theory of Change. This will provide a more robust framework against which the utility of the function can more comprehensively be assessed.

## 7 CONCLUSIONS AND RECOMMENDATIONS

251. The purpose of the Peer Review was to support UNICEF in identifying the steps required to further strengthen the function so that it is 'fully-fit' for purpose and well-placed to make the best contribution to the work and strategic positioning of the organization, whilst furthering developments in the field of evaluation. The core question that guided the Peer Review Panel was: "Are the agency's evaluation policy, function and its products: independent; credible; useful and influential for learning and accountability purposes, as assessed by a Panel of professional evaluation peers against the UNEG Norms and Standards and the evidence base?"
252. On the basis of the evidence available, analysed and discussed in the report, the Peer Review concluded that there is no short answer to this question, because the evaluation function in UNICEF aims at responding and being 'fit for purpose' for two deeply different views, of what is evaluation.
253. On the one hand, at the corporate level, UNICEF understands evaluation as a shared and decentralized management tool, that mostly contributes to evidence-generation together with research, studies and reviews; and whose contribution to the oversight function is not a primary concern. This is made explicit through the REP and the accountability framework for the function, which is strongly embedded with planning and management roles, and whereby evaluation duties and tasks, including approval of evaluation plans and reports, are compatible with planning and management roles and responsibilities on the subject being evaluated. Furthermore, in this view, a professional evaluation background is not a necessary requirement for a number of evaluation tasks, from the overall leadership of the function to the planning and management of the DEF.
254. On the other hand, stand the UNEG Norms and Standards for evaluation in the UN system, which are also a reference for the international evaluation community. The pillars of the N&S are the independence of the evaluation function, that is a fundamental condition for its credibility and with the latter, a necessary element of utility, in addition to other characteristics. From this perspective, evaluation is a function, more often than not associated to the corporate oversight function with Audit, which contributes to both accountability and learning. To achieve this, evaluation duties and responsibilities have to be clearly separate from any planning and management roles. UNEG is pursuing the professionalization of evaluation, with a complete set of competences for evaluators, as well as for those commissioning and managing evaluations.
255. Reality is in fact slightly more nuanced than in the description of the two 'schools of thought', and the Peer Review found evidence at various levels in UNICEF, of genuine efforts aimed at bridging the differences between the two. These include initiatives to strengthen the independence of the function; significant efforts to enhance the quality of the evaluation products and the utility of evaluations; real commitment to make available sufficient resources to the function to adequately fulfil its mandate. At the same time, meeting the requirements and expectations of both visions is very challenging and the Peer Review considers that many of the weaknesses it identified in its analysis, stem from the tension between the two.
256. This tension could lead to an impasse, if UNICEF decided that current evaluations already meet the corporate requirements on evaluation and that only minor adjustments are required, without challenging the fundamental nature of the function. Or, the same tension can become the inspiring challenge that leads UNICEF to transform its evaluation system into an independent, credible and extremely useful function that supports the organization to better implement its mandate and achieve its goals, in full respect of its decentralized nature. This option would be more in line with the strong commitment of the organization to its primary and ultimate clients, the children, and would make of UNICEF an innovator also in the field of evaluation.

257. Trusting that UNICEF will rise to this challenge, the Peer Review developed six major recommendations, each supported by several possible actions, for consideration by the Executive Board, Senior Management, and the leadership of the evaluation function. These address the need for a new evaluation policy that strengthens the independence of the function, with a direct effect on its credibility; the internal governance of the function, its resources, the quality of the evaluation process and the management of the EO. Each recommendation is supported by a brief synthesis of the evidence in the report that led to its formulation.

258. The Peer Review is aware that six recommendations, and the suggested actions, may appear to be 'many'. Not all will be regarded as acceptable or feasible, for a variety of reasons. However, because evaluation functions are complex systems with many interacting and inter-related attributes and processes, the Panel is reluctant to rank or prioritize the recommendations, and the related actions. All appear necessary if the ultimate goal is to make the UNICEF evaluation function fully fit for purpose, and respond to the needs of the organization while being aligned with the international evaluation standards.

#### *An independent and decentralized UNICEF Evaluation Function*

259. The report discussed many of the strengths of the UNICEF Evaluation function. It also identified a number of areas where adjustments are considered necessary, to bring the evaluation function fully in line with the UN Norms and Standards for evaluation and to enhance its effectiveness, and ultimately utility, for UNICEF itself and for its primary clients, the children of the world.

260. At the root of the weaknesses, key issues identified include the conceptual underpinning of a 'shared evaluation function' embedded in the REP accountability framework; the dilution of roles and responsibilities in evaluation with other roles and functions; the constrained role assigned to the EO with regard to the internal governance of the function. In the view of the Peer Review, there is no incompatibility between aligning these fundamentals of the function with the international norms on independence and credibility, while adjusting to the decentralized structure of the organization.

261. Recommendation 1 aims at addressing these issues, by focusing on the preparation of a new UNICEF Evaluation Policy, to be prepared in stronger compliance with the UNEG Norms and Standards and through an inclusive process that will draw views from stakeholders at all levels. The new Policy should explicitly integrate, and/or be informed by as appropriate, the other recommendations and related actions formulated by the Peer Review.

**Recommendation 1:** A new policy for the UNICEF evaluation function should be developed, that takes full account of the decentralized structure of the organization and integrates all the requirements for the independence, credibility and professionalization of the function, in line with the 2016 UNEG Norms and Standards, and with the other recommendations and actions proposed by the Peer Review.

#### ***Specific complementary measures for the implementation of Recommendation 1 include:***

1.1 Establish dual reporting line for Regional Evaluation Advisers: a direct administrative reporting line to the Regional Director, who would also ensure financial resources for the position; and a technical reporting line to the Director of Evaluation in headquarters;

1.2 Develop adequate impartiality provisions and safeguards for the behavioural independence of staff with responsibility in planning, commissioning and managing evaluations;

1.3 Make publicly available on UNICEF's external web site, all Management Responses to evaluations;

1.4 Revamp the Global Evaluation Committee as a platform for substantive discussion between the EO and UNICEF Management on, among others: evaluation topics and planning also on strategic

management issues; emerging and compliance issues; sharing of key evaluation findings of corporate relevance;

1.5 Development of a Theory of Change for the UNICEF evaluation function;

1.6 Development of a Policy implementation strategy, in the form of an Executive Directive, for the future evaluation policy, to guide its operationalization.

### *The internal governance of the evaluation function*

262. The report identified weaknesses in the internal governance of the evaluation function, due to the limitations of the roles and responsibilities assigned to the EO in terms of coordination, monitoring and oversight of the function as a whole. The status-quo: hampers the opportunities for realising synergies that stronger evaluation coordination at regional and central level may provide; undermines the credibility of the evaluation planning process; leads to duplication of efforts and waste of resources; and restricts improved learning from evaluations across the organization.

263. Recommendation 2 should not be interpreted as undermining the ultimate responsibility of the Executive Director for the evaluation function, but rather as a clear delegation of authority to the Director of Evaluation who would directly report to the ED on all evaluation matters.

264. Elements of the evaluation function where the evidence available showed that internal governance should be more active include: implementation and oversight of the new Policy; monitoring and oversight of agreed KPIs, and revision of KPIs for evaluation coverage; evaluation planning; and National Evaluation Capacity Development.

**Recommendation 2:** UNICEF should assign to the EO, full responsibility for the internal governance of the evaluation function, and adequate resources for fulfilling this role.

### ***Specific complementary measures for the implementation of Recommendation 2 include:***

The EO should take the lead, with inputs from the Regional Evaluation Advisers and in consultation with other relevant stakeholders in the organization, on the following actions among others:

2.1 Revision of the Key Performance Indicators on evaluation coverage at country level and on the budget allocated to the evaluation function to simultaneously achieve; adequate coverage of UNICEF's work; more accountability for the function itself; and more flexibility to adjust to the very diverse contexts and circumstances of UNICEF country and regional offices;

2.2 Ensuring that the Global Evaluation Plan includes strategic corporate-level evaluations, as well as evaluation syntheses or meta-evaluations, which draw on country and/or regional level evaluations identified in consultation with the Regional Evaluation Advisers;

2.3 Country level Costed Evaluation Plans should be discussed with both Regional Evaluation Advisers and the EO; these CEPs should include to the extent possible evaluations that will feed into regional, multi-country and global evaluations;

2.4 Regional Costed Evaluation Plans should be discussed with the EO and include as appropriate, evaluations that will feed into corporate-level evaluations.

2.5 Development of a corporate strategy for National Evaluation Capacity Development, taking into account the evolution of the debate on NECD within UNEG.

### *Financial resources for the evaluation function*

265. The report confirms the strong commitment of UNICEF to making adequate financial resources available to the evaluation function, and in pursuing the target of 1% of the resources, as foreseen by the REP. At the same time, the Peer Review identified potential risks to the credibility and impartiality of the function, due to the funding modality of a significant share of evaluation staff, in EO and in the DEF.

266. Evidence available suggests that the application of a 'blanket target' will not maximise efficiency in the allocation and use of available financial resources. The Peer Review considers that a more rational and efficient option would be elevating the goal of achieving the 1% target to the regional level. This would allow the development of evaluation plans of high strategic relevance to be implemented through regional and multi-country evaluations, as currently done in the CEE/CIS region; while remaining sufficiently close to the national contexts to identify the need for useful and mandatory evaluations at country level, such as those in the Myanmar example discussed above. This business model should be supported through some re-distribution of resources across countries within the same region and this would require advocacy and negotiations with donors whose regulations allow this type of mechanism.

267. With regard to the funding modality for evaluation staff positions, at least two measures could be envisaged to ensure that the EO and the DEF maintain access to sufficient human resources to meet the evaluation requirements of the organization, while enhancing the independence and credibility of the function. First, in both EO and Regional Offices as many evaluation positions as possible should be charged to the Regular Resources. Second, for all other positions and for covering the costs of regional and multi-country evaluations a sustainable mechanism should be developed to pool resources from both headquarter-based programme divisions and country offices, to be managed by the EO and the Regional Directors, as appropriate and relevant.<sup>47</sup> In this manner, the contributions from both Regular Resources and Programme Resources would accrue to the corporate spending on evaluation, in a more transparent and visible manner.

268. All these issues are tackled by Recommendation 3 and the related specific measures below.

**Recommendation 3:** UNICEF should establish new modalities for the funding of the evaluation function at all levels.

***Specific complementary measures for the implementation of Recommendation 3 include:***

3.1 The allocation of 1% of the financial resources of the organization spent on evaluation should be a target at the regional level, to enable a flexible and more efficient use of resources;

3.2 All EO and Regional evaluation staff positions should be funded through Regular Resources, or Programme Resources transferred to the EO and to the Regional Directors, managed under their respective direct responsibility;

3.3 A sustainable pool funding mechanism should be developed, to leverage resources from headquarters-based programme divisions and from country-offices, to fund evaluation specialist positions that cannot be funded through Regular Resources, and for conducting evaluations at the regional and/or multi-country level;

3.4 The multi-country evaluation specialist model tested in Cambodia, Malaysia and Myanmar, should be replicated across groups of countries where this may prove appropriate and useful to overcome scarcity of resources for evaluation at country-level.

*Human resources for the evaluation function*

269. Taking in consideration the evidence available about the staffing profiles, the mix of evaluation tasks with other roles and responsibilities, and the reporting lines of staff responsible for managing evaluations in the DEF, the Peer Review considers that the independence, credibility and quality of the DEF are negatively affected and that mitigating measures should be introduced to safeguard the professionalism and impartiality of the EO. This should include provisions on the rules

---

<sup>47</sup> FAO Office of Evaluation established such a mechanism in 2013, see [http://www.fao.org/fileadmin/user\\_upload/oed/docs/Evaluation\\_Docs/May\\_2016/FAO\\_procedures\\_funding\\_project\\_evaluations\\_final.pdf](http://www.fao.org/fileadmin/user_upload/oed/docs/Evaluation_Docs/May_2016/FAO_procedures_funding_project_evaluations_final.pdf).

governing staff rotation in the evaluation function, and all should be part of a long-term human resource organizational plan for strengthening the function.

270. The Peer Review also considers that the evaluation function would be more effective and efficient by establishing a cohort of full-time evaluation advisers and specialists located at the regional level, or at the national level where the national programme size justifies it, or as multi-country evaluation specialists, who would operate with technical reporting lines to the EO, while closely liaising with the regional and country offices to ensure the relevance and usefulness of evaluations.

271. Finally, the Peer Review also identified gaps, in relation to the UNEG Norms and Standards, in the REP and in the practice of the process for the selection and appointment of the Director of the EO. At the same time, with a new role for the Office of internal governance of the function as recommended above, it would appear appropriate to modify accordingly the title of the Director, to strengthen the coherence and unity of the function.

272. Recommendation 4 and the additional provisions below, address these issues.

**Recommendation 4:** UNICEF should develop a Strategic Human Resource plan for the evaluation function. The Plan should also contain a section that defines the provisions for the selection and appointment of the Director of the Evaluation Function

***The Strategic Human Resource plan for the evaluation function should include the following provisions:***

- 4.1 Establishment of the position of an EO deputy director, at D-1 level;
- 4.2 Ensuring that the Director of Evaluation and Deputy Director together provide credible evaluation experience and competence as required by UNEG Norms and Standards and by the UNEG Competencies Framework for UNEG Heads;
- 4.3 Establishment of the position of Regional Evaluation Adviser at the P-5 level in each and all UNICEF Regional Offices;
- 4.4 Through gradual reallocation of resources over time, establishment of teams of full-time evaluation advisers and specialists in each region, at regional, multi-country and national level where justified, by consolidating the resources currently used for M&E specialists at country level;
- 4.5 Development of specific provisions for the rotation of evaluation staff in UNICEF that allow staff to pursue a career in evaluation whilst maintaining their behavioural independence.

***The Strategic Human Resource plan for the evaluation function should include the following provisions for the selection and appointment of the Director of the UNICEF Evaluation Function***

- 4.6 The title of the position should become 'UNICEF Evaluation Director';
- 4.7 The Director of the Evaluation Function should be selected and appointed in agreement between the Executive Director and the Executive Board;
- 4.8 External evaluation expertise should be part of the selection panel for the Director, e.g. at the level of UNEG heads;
- 4.9 The Evaluation Director should report directly to the UNICEF Executive Director, on all matters;
- 4.10 The ToR for the Director of UNICEF Evaluation should include the systematic presentation of all global evaluation reports to the Executive Board, in addition to the Global Evaluation Plan and the Annual reports on the evaluation function as already the case.

### Quality of evaluations in UNICEF

273. UNICEF has significantly invested time and resources over the years, to enhance the quality of its evaluation products. This has resulted in a steady improvement of the quality of evaluation reports, as reported in the annual meta-synthesis based on the quality assessment of each evaluation report issued by the organization.
274. This notwithstanding, some challenges remain. More standardized evaluation processes and more guidance thereon are necessary, together with the need to expand the range of methods and tools applied, to improve the integration of cross-cutting issues into evaluations, and to enhance the quality and usefulness of recommendations. The suggested action to establish an External Evaluation Advisory Committee stems from the observation that such a body could contribute to support and enhance the quality of UNICEF evaluation strategies and products, at a time when the EO will go through important changes in staffing, mandate and responsibilities.
275. Recommendation 5 addresses these issues, focusing on the leadership role of EO in this regard.

**Recommendation 5:** All evaluations planned and commissioned by UNICEF, whether by EO, programme divisions, regional and country offices should aim at achieving the same standards of independence, credibility and utility, and align with the guidance and procedures established by the EO. The EO should be given the responsibility to strengthen the quality of the evaluation process in UNICEF, with an active oversight role on the DEF in this respect.

#### ***Specific complementary measures for the implementation of Recommendation 5 include:***

- 5.1 Establishment of an external evaluation advisory committee, similar to the one established by the Independent Evaluation Office in UNDP;
- 5.2 Development of standard guidance for quality assurance of the evaluation process;
- 5.3 Establishment of a real-time Quality Assurance mechanism in support of evaluations at country level in all regional offices, under the responsibility and management of the Regional Evaluation Adviser and the oversight of the EO;
- 5.4 Development and/or adaptation of evaluation guidelines and manuals to fit the UNICEF evaluation process, from inception to completion;
- 5.5 Development and/or adaptation of evaluation guidelines and manuals to improve the integration of human rights and gender equality perspectives in evaluations;
- 5.6 Revision of the criteria that differentiate evaluations from other types of assessments and reviews, and consider a more consistent application of the taxonomy in the titles of evaluation reports.
- 5.7 Revision, in consultation with Senior Management and the Regional Evaluation Advisers of the quality standards for evaluation recommendations, to make them more useful to UNICEF managers while maintaining the possibility for strategic, corporate and cross-cutting issues to be captured;
- 5.8 Revision of the timeframes for the implementation and closure of recommendations that address strategic and corporate-wide issues.

### Management of the EO

276. The Peer Review analysed different aspects of the management of the Evaluation Office, ranging from the internal arrangements for the management of corporate evaluations, to the work done to establish and manage evaluation information systems to meet its monitoring and reporting duties, to its efforts on National Evaluation Capacity Development.
277. The Office largely complies with its role and mandate. At the same time, areas for improvement were identified, in particular on the internal working arrangements to ensure the timely

delivery of evaluation reports and to enable a diversity of perspectives in evaluation. Recommendation 6 below addresses these issues.

278. The Peer Review is also aware that more changes will be necessary, also in terms of staff capacity, to meet the challenges raised by the stronger internal evaluation governance role entailed in Recommendation 2, and by Recommendation 5 on quality of evaluations. In this, the Advisory Panel recommended above could play an important role in supporting the Director of Evaluation in transforming the EO into an office better fit for its future purpose.

**Recommendation 6:** The Director of UNICEF Evaluation should revise the internal management processes of the EO.

***Specific complementary measures for the implementation of Recommendation 6 include:***

6.1 Improving the efficiency of the corporate-level evaluation processes and the timely delivery of all evaluation products by the EO;

6.2 Rotation of the EO evaluation advisers and specialists for the management and conduct of evaluations, so as to enable diversity of perspectives and experience in evaluations of the same thematic areas and programmes;

6.3 Inclusion of highly reputed specialists in the subject matter of the evaluation, into the evaluation teams responsible for carrying out evaluations commissioned by EO and the DEF;

6.4 Ensuring that all criteria and Key Performance Indicators identified in the Evaluation Policy are adequately monitored and reported upon.

**The future of the evaluation function**

279. UNICEF, even within the United Nations, is uniquely placed in terms of (i) the extensive global reach it has in 192 countries through country programmes and National Committees, seven regional offices, a supply operation in Copenhagen and offices in Geneva, Tokyo and Brussels, (ii) its highly decentralized operations, (iii) the presence of evaluation staff in many country offices, (iv) strong research capacity and a research centre in Florence, and (v) sizeable mobilization of resources from the private sector. In the broader development world, UNICEF has played a high profile and critical role in advocating for the rights of children and in producing evidence and data used by global and regional organizations. Looking ahead, we feel that:

280. UNICEF will need to capitalize much more strategically on the comparative advantages it enjoys. An independent UNICEF evaluation function in particular needs to keep pace with a rapidly changing development and evaluation world which, for its sustainability and progress, requires much greater equity, better assurance of rights, and more respect for human dignity.

281. The future of development will have to, in addition to ensuring effective service delivery, increasingly rely on generating and managing evidence and knowledge to scale effective interventions and on influencing a broad range of players and investors. The significance of financial contributions from UNICEF are likely to decrease even further as nations become richer and graduate to becoming middle or high-income countries. Scale partners such as governments and impact investors will depend on the knowledge that independent and credible evaluations, and research, can provide - of what works, when and where. This has several implications for UNICEF's evaluation function.

***Reputation and Influence***

282. Externally, the UNICEF EO continues to be well regarded and appreciated by the international evaluation community for its long-standing commitment to regional and national evaluation capacity development, and for the unique leadership role and contributions it has made to the global

development evaluation architecture (e.g. EvalPartners, AfrEA, SDGEval, tools on equity focused evaluation, webinars, MyMandE).

283. The evaluation community looks to UNICEF to continue to play this important role in an increasingly challenging global context for evidence and evaluation, and to be a much-needed champion for evidence of why equity and inclusion matter. These efforts will have to be more systematically assessed, streamlined and strengthened for greater impact on the lives of children.

### ***Strategic leadership***

284. The EO staff and its Director appear to have been skilful in navigating UNICEF's complex, decentralized system, building trust and relationships to engage in evaluation.

285. The EO has a window of opportunity now with its new Director of Evaluation to strengthen its leadership role in UNICEF's evaluation function and strategy by optimizing the evaluation synergies across the organization, evolving and responding to the need for more rapid cycle, strategic evaluations, synthesis reviews and cross organizational learning and accountability in addition to the ongoing evaluations requested of the EO.

### ***Evolving skills and capacities***

286. With strong leadership, there is also a window of opportunity for UNICEF's evaluation function to diversify its technical capacities to embrace evolving evaluation approaches and to be more influential with a wider range of key traditional and non-traditional actors in a post 2015 SDG and innovation era (from development policy players, investors, and private sector players). Three areas are likely to call for greater attention.

- d) UNICEF's evaluation function will need to equip itself better to deal with new technologies for real time data gathering and analyses through the use of new ICT technologies including the use of mobile, big data, predictive analytics, and machine reading technologies;
- e) UNICEF's evaluation work will need to encourage innovations in methodologies beyond traditional approaches to embrace complexity and systems approaches, stronger focus on rigorous social return on investment analyses (economic modelling of financial and social returns), and greater attention to efficiency and cost effectiveness;
- f) UNICEF evaluation staff will need to forge important forward looking innovation partnerships with, for example, the UNICEF Innovation Office and other evaluation Innovation Hubs to develop new, faster and more efficient approaches to monitoring and evaluation in UNICEF, the UN system and beyond.

### ***Learning and influence across complex systems***

287. Looking ahead (and even now), successful development outcomes rely on the actions of many diverse players and investors. No longer can one agency alone achieve global impact. The ability to learn across organizations, new partnerships, and actors will be crucial, and thus UNICEF's evaluation function will need to equip itself with a broader suite of tools and approaches to learn across systems, new players and to synthesize and communicate learning for influence.

288. UNICEF is likely to be called upon, given its large country-level and regional physical presence, to play a much more strategic leadership role in managing and learning from such partnerships.

289. Evaluations will be required to increasingly contribute to improving institutional performance, enhancing effectiveness and efficiency of collective actions, improve joint accountability, and utilize common learnings for delivering results.

290. The EO will have to play a bigger role in drawing on the quality, independence and coverage of the evaluations it conducts to draw lessons from meta-evaluations and cross-organizational learnings. In this regard, four areas where evaluation may provide insights are likely to warrant special attention:

- e) More can be done to optimize learning and strategic influence and to achieve greater synergies among the key evaluation players across the organization and the field;
- f) More needs to be known about the best ways for including and involving children (particularly adolescents) in evaluation. Though UNICEF and other organizations have promoted children's participation, and this is currently being pursued, it remains a neglected area;
- g) Better understanding is needed of combined impacts of 'development' and 'humanitarian' assistance on children and their families disrupted, upended or nearly destroyed by conflict, human-made disasters and wide-ranging, climate-change events and other emergencies;
- h) Ending violence is critical for sustaining development. Little however is known on prevention strategies that work to end violence particularly against children.

***Better positioning and synergies for greater influence***

291. The potential and opportunity exist for UNICEF to better use the Evaluation Function for strategic corporate evaluations and partnerships that look beyond programme sectors and sections to support effective corporate governance, greater synthesis of knowledge and lessons (with research), risk management (with audit), and strategy and leverage of private sector lessons and investments. One example, among many possible topics, would be to use findings from an evaluation of the private sector's contribution to UNICEF in helping to mobilize greater resources and attention to improving the lives of children.