

UNEG Peer
Review Series



UNEG
United Nations Evaluation Group

UNEG PEER REVIEW OF THE EVALUATION FUNCTION OF UNICEF

April 2023

UNEG/REF(2023)

EXECUTIVE SUMMARY	V
PURPOSE AND TERMS OF REFERENCE OF PEER REVIEW	V
DATA COLLECTION	V
THE PEER REVIEW PANEL.....	V
KEY CONCLUSIONS.....	VI
<i>Independence</i>	<i>vi</i>
<i>Credibility</i>	<i>vi</i>
<i>Utility</i>	<i>vii</i>
<i>Roles and Responsibilities</i>	<i>vii</i>
<i>Follow-up and use of evaluations</i>	<i>vii</i>
<i>The enabling environment</i>	<i>viii</i>
RECOMMENDATIONS	VIII
RECOMMENDATION 1.....	VIII
RECOMMENDATION 2.....	VIII
RECOMMENDATION 3.....	VIII
RECOMMENDATION 4.....	IX
RECOMMENDATION 5.....	IX
INTRODUCTION TO THIS REPORT	1
BACKGROUND TO THE PEER REVIEW	1
TERMS OF REFERENCE	1
THE 2018 EVALUATION POLICY OF UNICEF	1
THE UNICEF EVALUATION FUNCTION IS SUPPORTED BY A FORMAL INFRASTRUCTURE OF GOVERNANCE, ROLES AND RESPONSIBILITIES, AND POLICIES, GUIDANCE AND TOOLS	5
SITUATING THE UNICEF EVALUATION SYSTEM WITHIN ITS BROADER CONTEXT	6
THE EVALUATION CHALLENGES FACED BY UNICEF ARE NOT UNIQUE AMONG ORGANIZATIONS	8
INTRODUCTION	10
INDEPENDENCE	10
DEFINITION OF INDEPENDENCE.....	11
THE PRINCIPLE OF INDEPENDENT EVALUATION ENJOYS WIDE SUPPORT WITHIN UNICEF	11
INDEPENDENCE SHOULD BE FURTHER OPTIMIZED BY ADDRESSING CERTAIN STRUCTURAL AND BEHAVIOURAL FACTORS	12
PURSUING SOME MORE EMBEDDED AND MORE COLLABORATIVE, SYNERGISTIC APPROACHES, WHERE RELEVANT AND APPROPRIATE, COULD HELP TO IMPROVE RELEVANCE AND UTILITY	13
CREDIBILITY	14
DEFINITION OF CREDIBILITY.....	14
THERE HAVE BEEN MEASURABLE IMPROVEMENTS IN EVALUATION QUALITY	15
FINANCIAL AND HUMAN RESOURCES STRENGTHEN THE CREDIBILITY OF THE EVALUATION FUNCTION	19
FURTHER PROGRESS ON CREDIBILITY IS LIMITED BY A RANGE OF FACTORS	19
<i>Improving consistency in decentralized evaluation quality and the appropriateness of methodological approaches</i>	19
<i>Harnessing the evaluation portfolio to strategic priorities and to what matters the most</i>	21
<i>Improving the availability of monitoring data</i>	22
<i>Ensuring that funding modalities support effective evaluation for accountability and learning</i>	22
UTILITY	23
DEFINITION OF UTILITY	24
VIEWS ON THE UTILITY OF EVALUATIONS ARE MIXED	24
<i>Improving selection of evaluation topics</i>	25
<i>Improving the planning and management of evaluations, including through strengthening stakeholder engagement</i>	26

<i>Ensuring the timeliness of evaluations</i>	27
<i>Improving the format and accessibility of evaluation products</i>	27
FORMAL COMPLIANCE WITH MANAGEMENT RESPONSE PROCESS.....	28
ROLES AND RESPONSIBILITIES	28
FORMAL ROLES AND RESPONSIBILITIES	28
ACCOUNTABILITY ARRANGEMENTS AT HQ AND DECENTRALIZED LEVELS	28
ROLES, RESPONSIBILITIES AND ACCOUNTABILITIES CARRY RISKS TO INDEPENDENCE	29
MECHANISMS FOR ENSURING ACCOUNTABILITY IN RELATION TO FUNDING, ESTABLISHING ADEQUATE CAPACITY AND APPLYING NORMS AND STANDARDS COULD BE STRENGTHENED	30
POTENTIAL OPTIONS FOR FORMAL ROLES AND RESPONSIBILITIES.....	31
FOLLOW-UP ON EVALUATIONS AND USE	32
VERY HIGH FORMAL COMPLIANCE WITH MANAGEMENT RESPONSE PROCESSES DOES NOT NECESSARILY INDICATE THAT EVALUATIONS ARE WELL USED WITHIN DECISION-MAKING.....	32
THERE IS POTENTIAL FOR GREATER COLLABORATION ACROSS UNICEF KNOWLEDGE FUNCTIONS TO IMPROVE THE WIDER VISIBILITY, AND SO SUPPORT THE USE, OF EVALUATIONS.....	33
THE ENABLING ENVIRONMENT.....	34
<i>Defining 1 per cent</i>	35
<i>Modalities of funding</i>	35
<i>Is 1 per cent sufficient?</i>	36
STRENGTHENING UNICEF'S EVALUATION CULTURE OUTSIDE EVALUATION PRACTITIONERS	37
MAXIMIZING THE BENEFITS OF EVALUATION THROUGH COLLABORATION WITH OTHER FUNCTIONS	38
LINKING EVALUATION WITH DIGITAL TRANSFORMATION.....	38
THE EXTERNAL ENABLING ENVIRONMENT FOR EVALUATION	39
RECOMMENDATIONS	44
RECOMMENDATION 1.....	44
RECOMMENDATION 2.....	45
RECOMMENDATION 3.....	45
RECOMMENDATION 4.....	46
RECOMMENDATION 5.....	47

Acronyms and abbreviations

CEP	Costed Evaluation Plan
CO	UNICEF Country Office
CPD	Country programme document
DAPM	UNICEF Division of Data, Analytics, Planning and Monitoring
DEF	Decentralized Evaluation Function
EA	Evaluability Assessment
EB	Executive Board
ED	Executive Director
EISI	Evidence Information Systems Integration database
EMOPS	UNICEF Office of Emergency Programmes
EMR	Evaluation Management Response
EO	UNICEF Evaluation Office
EQA	Evaluation Quality Assessment
EvalNet	OECD-DAC Network on Development Evaluation
G staff	General Service staff (G1 the most junior, G7 the most senior)
GEROS	Global Evaluation Reports Oversight System
HQ	UNICEF headquarters in New York
IATI	International Aid Transparency Initiative
ICT	Information and Communication Technology
JIU	Joint Inspection Unit
KPI/s	Key Performance Indicator/s
M&E	Monitoring and Evaluation
MOPAN	Multilateral Organisation Performance Assessment Network
MPTF	Multi-Partner Trust Funds
NECD	National Evaluation Capacity Development
ODIL	Organizational Design and Improvement Lab
OECD-DAC	The Development Assistance Committee of the Organisation for Economic Co-operation and Development
P staff	Professional staff level (P1 the most junior, P6 the most senior)
REA	Regional Evaluation Adviser
SDGs	Sustainable Development Goals
SOP	Standard Operating Procedures
SP	Strategic Plan
ToR	Terms of Reference
UNDS	United Nations Development System
UNEG	United Nations Evaluation Group
UNEG N&S	United Nations Evaluation Group Norms and Standards
UNDP	United Nations Development Programme
UNGA	United Nations General Assembly
UNICEF	United Nations Children's Fund

Executive summary

Purpose and Terms of Reference of Peer Review

In common with all UNEG Peer Reviews, this report aims to identify existing good practices and to strengthen further the UNICEF evaluation function. The approach reflects UNEG's *Norms and Standards* and seeks to answer the strategic question posed in the Peer Review Terms of Reference:

Based on the experience implementing the 2018 Evaluation Policy for UNICEF over the past five years, what aspects of the policy have worked well in practice in optimizing the evaluation function in the organization in accordance with the UNEG norms and standards and should therefore be retained and potentially built on, which aspects have not worked well and should be changed or abandoned, and, in retrospect, what gaps evident in the 2018 policy need to be filled in the 2023 Evaluation Policy?

UNEG Norms and Standards identify six areas to be addressed: independence; credibility; utility; roles and responsibilities; use of, and follow up on, evaluation; and the enabling environment.

Data collection

The report draws upon an extensive range of data collection methods, including:

- Secondary analysis of UNICEF datasets, including the results of UNICEF's self-assessment exercise conducted with 241 evaluation practitioners and 171 senior management staff
- Review of 44 documents, including policies, Standard Operating Procedures and reports
- Detailed information made available through the Evaluation Office
- Twenty-eight interviews with UNICEF stakeholders
- Five Focus Group Discussions with UNICEF stakeholders
- Nine interviews with external stakeholders (including from programme and donor governments, external evaluation consultancies and other partners)
- Observation and participation at the Global Evaluation meeting, 7-11 November 2022
- A light-touch review of evaluation functions within four comparator organizations
- An online survey of Executive Board members
- Consultation with UNICEF's executive board, through workshops with regional groups

The peer review panel

The panel comprised:

- Lisa Sutton, Director, Independent Evaluation and Audit Services, UN Women (Panel Chair)
- Alan Fox, Deputy Director, Independent Evaluation Office, UNDP
- Jo Kaybryn, Head of Evaluation Office, International Committee of the Red Cross
- Dr. Peter van der Knaap, Director of the Policy and Operations Department of the Ministry of Foreign Affairs of The Netherlands
- Dr. Winston J. Allen, Agency Evaluation Officer, Bureau for Policy, Planning and Learning, United States Agency for International Development (USAID)
- Dr. Zakari Lawal, the National Director of Monitoring and Evaluation at the Federal Ministry of Finances, Budget and National Planning of Nigeria

Two consultants were appointed to support the data collection and analysis of the review panel:

- Dr. Tom Ling, Head of Evaluation, RAND Europe and President of the European Evaluation Society
- Dr. Ana FitzSimons, Research Leader, RAND Europe

Key Conclusions

UNICEF's progress towards delivering the 2018 *Evaluation Policy* aims is broadly good, with some areas for enhancement. Policies, guidelines, data on evaluation quality, management responses and the existence of a specific funding target for evaluations, are areas of strength in UNICEF's evaluation function. The number of evaluations has doubled since 2018. The function can produce high-quality, useful and credible evaluations.

The limitations of the evaluation function centre around the wider enabling environment and ensuring accountability and monitoring for roles and responsibilities, including funding and reporting lines, which can limit consistent achievement of independence, credibility and utility. Further risks concern the limitations to the management use of evaluations and the interactions between the evaluation function and other knowledge functions within UNICEF.

The Evaluation Office has responsibilities for addressing some of the limitations but, for its efforts to have maximal effect, also needs to be empowered and supported by other elements within the enabling environment, including UNICEF leadership and management in other functions.

Independence

- The principle of maintaining the independence of the evaluation function has high visibility in evaluation guidance and policy documents and enjoys strong support among UNICEF stakeholders. The UNICEF Executive Board, like the Executive Boards of other UN entities, is placing increased attention on ensuring the independence of oversight functions.
- At the global level, the independence of the function has been enhanced since 2018, including by appointing senior staff with substantial evaluation expertise and by introducing in the policy a specific maximum term for the Director of Evaluation's service.
- Key issues affecting the independence of the evaluation function, particularly at the decentralized level, were identified. These include funding modalities and the formal structure of roles and accountabilities, which neither strongly enforce nor incentivize behavioural independence. This results in variation in the level of independence in decentralized settings.
- While the independence of evaluations is rightly valued as among the key enablers of credibility and utility and as crucial to oversight and accountability, as the operating environment of UNICEF evolves and becomes more complex, there will be times when pursuing more embedded and collaborative approaches to the design and conduct of evaluations may contribute towards strengthening the relevance and utility of findings and recommendations.

Credibility

- Geros data suggest a strong and improving situation in relation to the quality and geographical coverage of UNICEF evaluations and to incremental improvements to the integration of disability and gender equality considerations within evaluations. Geros itself has emerged as a key enabler of quality and credibility.
- Most senior managers and evaluation practitioners in UNICEF regard the UNICEF evaluation function as credible, though there is room for improvement.
- The credibility of the evaluation function has been strengthened since 2018 by investments in human and financial resources.
- Progress on credibility is affected by a range of factors, some of which are within the control of the evaluation function and some of which would require UNICEF management action. Factors of the first category include improving the consistency in the quality of decentralized evaluations

and their methodological approaches and better focusing the portfolio of evaluations on UNICEF's overarching strategic priorities and on areas where evidence is lacking or most needed. UNICEF management action is required to ensure both the availability of high-quality monitoring data and effective funding modalities that support evaluation planning and implementation as well as application of learning within workplans.

Utility

- The utility of evaluations is a result of a relationship between demand (interest in, awareness of and capacity to use recommendations) and supply (the communication of relevant, actionable and evidence-based findings and recommendations). Not all of this is within the control of the evaluation function.
- Among senior management survey respondents, 76 per cent felt evaluative exercises had been useful to their office or division.
- The review found some important barriers to utility along the evaluation process from conceptualizing the issues to be evaluated, to recruiting evaluation consultants, to communicating evaluation results.
- One way that utility could be strengthened would be by including more voices, including those of children and youth, when scoping the evaluation issues and identifying recommendations as well as ensuring the productions of evaluation products in easily accessible formats.
- Practitioners expressed concerns about the quality of available external providers willing to work for the rates available, who may not sufficiently understand UNICEF's needs and ways of working.
- Management responses to evaluation are nearly always produced, which is positive, but a gap remains in evidencing that evaluations have been used to improve policy, programming, practice and outcomes for children.

Roles and Responsibilities

- While most roles, responsibilities and accountabilities for evaluation were clearly set out in the *Revised Evaluation Policy*, there are some areas that could be strengthened at all levels by clarification in the coming policy revision.
- Clarity of roles and responsibilities for evaluation neither consistently translates into independence of the evaluation function, particularly at decentralized levels, nor supports the wider aims of the *Revised Evaluation Policy*.
- Mechanisms for enforcing accountability for evaluation, including reporting and monitoring, currently result in variation in how roles are fulfilled throughout the evaluation function, particularly at decentralized levels, and should be strengthened.
- Alternative potential options for strengthening accountability for evaluation each carry their own potential advantages and challenges.

Follow-up and use of evaluations

- There is a high level of formal compliance with management response processes, but this does not necessarily indicate that evaluations are well used by the organization as a whole or by individual programmes.
- The Evaluation Management Response process is helpful for monitoring purposes but could be routinely supplemented by engagement with evaluation users to achieve a deeper understanding of what they value most highly and what could be improved.
- There is potential to create a more user-friendly digital repository of evidence and to promote this among UNICEF and wider UN staff as well as among external policymakers and researchers, which would in turn help to secure UNICEF's position as leaders in generating and organizing evidence to improve the situation of children.

- There is potential for improvement to the evaluation communications function to craft effective messages communicating the evaluation findings and implications for target audiences in order to increase the visibility and to enable broader use of findings.

The enabling environment

- Modalities of funding, including the irregularity of when and how much funding is allocated to evaluation at different levels, do not currently consistently support effective planning and commissioning of evaluations. It is management's responsibility to ensure that sufficient and predictable funds are provided for evaluation.
- The sufficiency of the 1 per cent allocation for evaluation was challenging to answer with the available evidence. The peer review learned of concerns regarding its calculation.
- The theory of change for evaluation in UNICEF is welcome but could be better used as a living document and could be further developed, with greater attention to what is within the evaluation functions' sphere of control and what is within its sphere of influence.
- UNICEF is driven by laudable goals and the evaluation function struggles at times for organizational attention in this environment. Strengthening UNICEF's evaluation culture outside the evaluation function would help to ensure that evidence on what works is valued as fundamental to the mission of improving the situation of children and is applied within management decision-making. This requires action not only from the evaluation function but also from UNICEF leadership.
- More could be done to maximize the benefits of evaluation through collaboration with other functions (including ethics, data, innovation and communication functions), which is not currently routine. We support the Director of Evaluation's strategic vision in this regard.
- In particular, linking evaluation with digital transformation is a substantial challenge for the future, but it is essential to ensuring that UNICEF can leverage the advantages of digitally enabled evaluation methods.

Recommendations

The recommendations listed here reflect the recognition that UNICEF is continuing to make good progress towards strengthening its evaluation function to support UNICEF's mandate to advocate for the protection of children's rights, to help meet their basic needs and to expand their opportunities to reach their full potential. Therefore, many recommendations seek to continue or amplify existing approaches. Other areas where new challenges and new opportunities are apparent may require a change of direction. Suggestions for complementary measures to advance the recommendations are included in the report.

Recommendation 1

The main aims of the 2018 Revised Evaluation Policy to support independence, credibility and usefulness through centralized and decentralized evaluation structures should be continued. There is room for further strengthening of roles, responsibilities and accountabilities, particularly at the decentralized level.

Recommendation 2

UNICEF management should ensure the sufficient and predictable funding of the evaluation function. The 1 per cent target should be clarified, including how it is calculated and what is included. Governance and criteria for allocation and use of the Evaluation Pooled Fund should be clarified.

Recommendation 3

The evaluation function should build on its added value by identifying, together with other knowledge functions in the organization, specific and focused ways to cooperate more closely.

Recommendation 4

UNICEF should maintain and further strengthen the improvements that have been achieved in its evaluation of human resources and leadership capacities.

Recommendation 5

UNICEF should use its mandate and visibility to work with and influence the external evaluation environment of UN agencies, academia, voluntary evaluation organizations and foundations to promote national evaluation capacity development and to mobilize evaluation action in support of UNICEF's mission.

Introduction to this report

Background to the peer review

The United Nations Evaluation Group (UNEG) has, since 2004, supported professional Peer Reviews of the evaluation functions of UN organizations. Organizations participate voluntarily. Each Peer Review should identify good practice and opportunities to further strengthen the evaluation function in the agency under review. Peer reviews should take into account the UNEG Norms and Standards and UNEG Code of Conduct.

Terms of reference

Full terms of reference for this peer review are available in Annex 1. This peer review was conducted under the auspices of UNEG and answers the following strategic questions:

Based on the experience implementing the 2018 Revised Evaluation Policy for UNICEF over the past five years, what aspects of the policy have worked well in practice in optimizing the evaluation function in the organization in accordance with the UNEG norms and standards and should therefore be retained and potentially built on, which aspects have not worked well and should be changed or abandoned and, in retrospect, what gaps evident in the 2018 policy need to be filled in the 2023 Evaluation Policy?

The peer review was completed in the first quarter of 2023 and should inform the evaluation policy revision that is planned to be completed in June 2023 and endorsed by the Executive Board in September 2023. The review also offers observations with the aim of strengthening the evaluation function more broadly. To this end, the review provides a systematic, impartial and independent evidence-based assessment of the UNICEF evaluation function over the past five years and reflects on how well and how far the 2018 *Evaluation Policy for UNICEF* has been implemented. It is a forward-looking exercise offering concrete recommendations to further strengthen UNICEF's evaluation function during the UNICEF Strategic Plan 2022-2025 and beyond.

The key audiences for the review include the UNICEF Executive Board, UNICEF Executive Director and Deputy Executive Directors and the wider management of UNICEF, the UNICEF Global Evaluation Committee, the UNICEF Director of Evaluation and those staffing the evaluation function at all levels of the organization. The review will also be of interest to a wider audience and will be available on the UNEG website.

The 2018 Evaluation Policy of UNICEF

A first UNEG peer review of the UNICEF evaluation function was conducted in 2006, with a subsequent evaluation policy approved by the UNICEF Executive Board in 2008. In 2013, following another review of the function, the Executive Board approved a revised evaluation policy. In 2017, a further peer review of the evaluation function was conducted and, in July 2017, its report was published, with the Management Response published in December 2017. The UNICEF *Evaluation Policy* was revised again in June 2018 in accordance with Executive Board Decision 2018/2.

The 2017 peer review¹ identified several strengths in the evaluation function, including commitment to and progress towards allocating 1 per cent of its financial resources to evaluation, improvements in evaluation quality over time and concurrent improvements in compliance with requirements for Evaluation Management Responses. It also identified several ways in which the function could be strengthened and developed six overarching recommendations to help institutionalize a more

¹ <http://www.uneval.org/document/detail/2045> (last accessed 25 May 2022)

independent and decentralized UNICEF evaluation function. These recommendations are set out in Annex 2 of this report. The recommendations informed the current institutional arrangements of the evaluation function as articulated in the 2018 *Evaluation Policy of UNICEF*² and the *UNICEF Procedure on the Implementation of the 2018 UNICEF Evaluation Policy*.³ The policy embodied numerous significant recommendations from the review, including a strengthened capacity for independent and decentralized evaluation (such as the creation of the post of Regional Evaluation Adviser in each of the organization's seven regional offices), clearer and more predictable resourcing of the function and more.

Brief context since 2018

Since June 2018, much has changed in the operating environment of UNICEF. COVID-19 has had an impact on the work of UNICEF and, at the same time, progress towards Sustainable Development Goals (SDGs) has stalled. There has also been a substantial increase in UNICEF's budget and a heightened interest shown by the international community in UN agencies' accountability, organizational learning and capacity to deliver at ever-greater scale. However, the key norms and standards, identified by UNEG, remain relevant and include independence, credibility, utility, impartiality, ethics, transparency, human rights and gender equality, national evaluation capacities and professionalism. Some of the challenges facing UNICEF were well summarized in the 2020 MOPAN assessment report:⁴

UNICEF's mission on behalf of the world's children faced significant external challenges: sustainable development gaps, persistent humanitarian crises and fragility, the effects of the COVID-19 pandemic, and an environmental crisis and climate change. At the same time, it faced the internal challenges of translating programme results into more significant gains at the outcome level, working with increasingly high levels of earmarked funding and insufficient core resources, implementing new system-wide changes required by United Nations Development System (UNDS) reform, doing more to combat sexual exploitation and abuse and sexual harassment, and capitalising on opportunities associated with innovation, shared services, and digitisation.

Peer review scope, methodology and limitations

The scope is taken from the Terms of Reference for the peer review. The peer review is concerned with the entire UNICEF evaluation function, including at the headquarters, regional and country levels. In line with the Terms of Reference, the peer review analyses and provides a diagnostic of key stakeholders' overall perceptions of six main foci: evaluation independence, credibility and utility, roles and responsibilities for evaluation, use of and follow-up on evaluations, and the overall enabling environment for the evaluation function.

Our methods included:

- Secondary analysis of UNICEF datasets, including the results of UNICEF's self-assessment exercise conducted with 241 evaluation practitioners and 171 senior management staff
- Review of 44 documents, including policies, Standard Operating Procedures, and reports

² Revised evaluation policy of UNICEF, 2018. <https://www.unicef.org/media/54816/file>

³ UNICEF Procedure on the implementation of the 2018 UNICEF Evaluation Policy. Document Number: EVAL/PROCEDURE/2018/001 Effective Date: 18 October 2018.

⁴

[https://www.mopanonline.org/assessments/unicef2020/MOPAN%20Assessment%20UNICEF%20web%20for%20download\].pdf](https://www.mopanonline.org/assessments/unicef2020/MOPAN%20Assessment%20UNICEF%20web%20for%20download].pdf)

- Detailed information made available through the Evaluation Office
- Twenty-eight interviews with UNICEF stakeholders
- Five Focus Group Discussions with UNICEF stakeholders
- Nine interviews with external stakeholders (including from programme and donor governments, external evaluation consultancies and other partners)
- Observation and participation at the Global Evaluation meeting, 7-11 November 2022
- A light-touch review of evaluation functions within four comparator organizations
- An online survey that was sent to Executive Board members
- Consultation with UNICEF's executive board, through workshops with regional groups

Further information on documents and stakeholders consulted is available in Annexes 3 and 4. Data was triangulated and analysed thematically in accordance with an evaluation matrix structured by the six key foci of the peer review. The draft report was shared with the Evaluation Office, whose factual corrections, insights and suggestions were integrated as considered appropriate.

Limitations

The peer review of UNICEF's evaluation function was not intended to provide a full and formal evaluation of the function's practices, processes and outcomes. The main limitations of the review include resources and the compressed timeline in which the review took place, and the non-representative nature of the sample of stakeholders we consulted with. Overall, there was greater engagement in our consultations among stakeholders from UNICEF leadership than among country office personnel. Some data collection methods yielded limited results, such as the survey for Executive Board members, for which only one response was returned, and there were limited participants in the regional consultations. Given UNICEF's work in over 190 countries, the peer review was never intended to consult with a representative sample of all country staff, but the limited number of stakeholders from decentralized settings is nonetheless acknowledged as a limitation of the methods. In mitigation of the non-representative samples of stakeholders engaged specifically for this peer review, we conducted secondary analysis of UNICEF datasets, including the results of UNICEF's self-assessment exercise conducted with 241 evaluation practitioners and 171 senior management staff.

Peer review time period (2018-2023)

This peer review focuses on the period covered by the current evaluation policy but locates the analysis of this period within a wider time frame when appropriate. In particular, the review briefly looks back to earlier peer reviews and evaluation policies where this helps to explain and illuminate actions and consequences in 2018-2022. This covers a time in which there has been a growing recognition of the need to re-energize progress towards achieving the 2030 Agenda and, therefore, the role that UNICEF plays in this. The relationship between these future goals and current evaluation practices and policies will therefore also be considered. This is also a period during which COVID-19 impacted the focus of evaluations and methods of data collection. More recently, several global and regional crises have required attention, alongside longer-running humanitarian crises.

Introduction to the UNICEF evaluation function and its context

UNICEF's evaluation function produces diverse products with wide coverage

This peer review adopts the UNICEF view that evaluation is “an assessment, conducted as systematically and impartially as possible, of an activity, project, programme, strategy, policy, topic, theme, sector, operational area or institutional performance. It analyses the level of achievement of both expected and unexpected results by examining the results chain, processes, contextual factors and causality using such appropriate criteria as relevance, effectiveness, efficiency, impact and sustainability. An evaluation should provide credible, useful evidence-based information that enables the timely incorporation of its findings, recommendations and lessons into the decision-making processes of organizations and stakeholders” (UNICEF 2018, p. 22).

UNICEF views evaluation as one of “several distinct complementary functions, which include child rights monitoring, performance monitoring and broader results-based management, research and audit. These functions support an enabling environment for the conduct of evaluations and their use. Data from monitoring and research informs evaluation; methods from research complement those used for evaluation, and audit findings help to shape evaluations and contribute to assessments of organizational efficiency. While maintaining its independence, evaluation will coordinate with those complementary functions” (UNICEF 2018, p. 13).⁵ In UNICEF, these ‘complementary functions’ sometimes take place within parallel lanes, and we will discuss the opportunities that exist for closer collaboration.

The review recognizes the diverse and growing range of evaluation products in UNICEF, which has almost doubled since the 2018 Evaluation Policy. The annual report for 2021 on the evaluation function in UNICEF sets out the steadily rising numbers of evaluative products produced by the function: 107 in 2018, 116 in 2019, 155 in 2020 and 178 in 2021.⁶ In 2022, 199 evaluation reports were recorded.⁷

Some nine out of 10 of these evaluation products are now managed at the country or regional level within a decentralized evaluation system (similar to other UN agencies such as UNDP and WFP). In some instances, evaluations are commissioned and managed in partnership with development partners. This decentralized system is intended to ensure that the evaluation function is infused across an organization – not solely centralized and imposed ‘top-down’ – with the intention that country-level managers should be fully invested in the analysis of what works and what does not. If implemented effectively, a decentralized approach creates considerable opportunities (to test new ideas, provide relevant and real-time accountabilities and build an ecosystem that is evidence-rich and evidence-driven) but it also presents challenges (achieving a balanced and prioritized portfolio of evaluations that support delivery of an organization’s mandate and entrench accountability). Country and regional evaluations are delivered by external consultants and their work must be commissioned, managed and quality-assured and the knowledge gained must be mobilized to support accountability and learning.

Other independent evaluations are commissioned by the Evaluation Office at Head Office (HQ) level in line with the Plan for Global Evaluations 2022-2025⁸ and following consultation with the Global Evaluation Committee, Programme Division and others. For example, in October 2022, there was an

⁵ <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N18/106/37/PDF/N1810637.pdf?OpenElement>

⁶ Annual report for 2021 on the evaluation function in UNICEF, available [here](#)

⁷ [Evaluation reports | UNICEF Evaluation in UNICEF](#)

⁸ <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N21/387/24/PDF/N2138724.pdf?OpenElement>

Evaluability Assessment and Formative Evaluation of the UNICEF Positioning to Achieve the Goals of the Strategic Plan 2022-2025, which was managed by a small team from the UNICEF Evaluation Office, overseen by an Evaluation Reference Group (of internal stakeholders) and conducted by an external evaluation team. In addition, UNICEF commissions or takes part in several joint evaluations and syntheses, as summarized in Annex 5 of this report.

The anticipated foci of evaluation topics are identified in the UNICEF Strategic Plan 2022-2025, which highlights five Goal Areas and envisages that, in each Area, there should be 'Data, research, evaluation and knowledge management: generating research, data and other forms of evidence, and applying and sharing it to support policies, planning, partnerships, resource mobilization, decision-making and programme delivery'. The *Plan for Global Evaluations 2022-2025*⁹ closely reflects the Strategic Plan and the evaluation portfolio at region and country levels responds to the Strategic Plans as mediated by the formal allocation of tasks and responsibilities outlined in Annex 6. The policy also establishes evaluation coverage norms:¹⁰

1. Country programme evaluations: At least once every two programme cycles, sequenced to feed into subsequent CPD and UNDAF
2. Country office-level thematic evaluations: At least one country thematic evaluation, country programme component evaluation, or project evaluation per year for each country programme
3. Component evaluation or project evaluation per year for each country programme
4. Evaluation of humanitarian action: Various norms of L1, L2 and L3 emergencies
5. Corporate evaluations: Themes under each Strategic Plan Goal Area, cross-cutting priorities
6. Evaluability assessments, evaluation syntheses, meta-evaluations: Determined by commissioning office
7. Reviews in support of policies, plans and strategies: Determined by commissioning office
8. Regional-level multi-country evaluations: Determined by the regional office
9. Joint United Nations and system-wide evaluations including UNDAFs and joint programmes, and the Strategic Plan common chapter: Determined by interagency mechanism
10. Country-led evaluations: Determined by partner governments

The UNICEF evaluation function is supported by a formal infrastructure of governance, roles and responsibilities, and policies, guidance and tools

In terms of the governance of the UNICEF evaluation function, the *Evaluation Policy* is endorsed by the Executive Board, which also endorses the plan for global evaluations and approves the budget of the Evaluation Office. The Executive Director is responsible for safeguarding the integrity of the evaluation function and promoting a culture of learning and accountability throughout UNICEF. The Director of Evaluation is accountable for the oversight of the function and leads the implementation of the *Evaluation Policy*. An Audit Advisory Committee advises the Executive Director and includes external experts and is responsible for upholding norms and standards. An Evaluation Advisory Panel advises the Director of Evaluation and recommends improvements in evaluation approaches and methodologies. A Global Evaluation Committee is an internal group focused on using evaluation evidence.

⁹ https://www.unicef.org/executiveboard/media/8911/file/2022-3-Plan_for_global_evaluation-EN-ODS.pdf

¹⁰

[https://www.unicef.org/evaluation/media/1241/file/UNICEF%20Evaluation%20Policy%20Brief%20\(2018\).pdf](https://www.unicef.org/evaluation/media/1241/file/UNICEF%20Evaluation%20Policy%20Brief%20(2018).pdf)

The UNICEF evaluation function is highly decentralized: at Headquarters, division directors should support evaluation policy, practice and capacity; regional offices should ensure the timely production of evaluations and their application of norms and standards; and, in country offices, representatives should ensure that evaluations are resourced, credible and used (while also linking with country stakeholders such as government officials). The formal allocation of tasks and responsibilities for implementing UNICEF's evaluation function, as envisaged in the 2018 *Evaluation Policy*, is set out in Annex 6 of this report.

Delivery of the evaluation function is supported by a total of 31 personnel (including consultants) employed at Headquarters. The number of HQ evaluation personnel, by level and type of contract, as well as the gender balance HQ evaluation personnel, is set out in Annex 7.

Alongside the governance, management and staffing arrangements, UNICEF's evaluation system is supported by an extensive array of formal tools, guidance, policies, Standard Operating Procedures (SOPs) and practices set in place to ensure the commissioning, delivery and mobilization of evaluation products. Key tools and guidance are set out in Annex 8.

Situating the UNICEF evaluation system within its broader context

The governance, staffing and guidance of the evaluation function, discussed above and illustrated by the inner ring in Figure 1 below, are situated within a broader enabling environment for evaluation. There is, first, an institutional enabling environment where, it is intended, the 'distinct and complementary functions' create synergies, allowing the evaluation system to be informed by, and to inform, other UNICEF institutional functions, in pursuit of their shared mandate (programming, partnerships, advocacy and innovation). This is illustrated in the outer ring in Figure 1. There is also, second, an external enabling environment of scholars, think tanks, policymakers, donors and innovators in securing rights for children. UNICEF's evaluation system also contributes to, and draws upon, this wider world of thinking and practice. Some of this wider thinking is communicated in a series of webinars and e-Learning opportunities.

The evaluation function at UNICEF is well-regarded. In 2020, MOPAN noted that UNICEF's "evaluation function also scored well for its independence, tracking of poor performance and follow-up systems. It showed room for improvement on decentralized evaluations, including their coverage and quality, as well as for performance monitoring, including having evidence-based targets, effective monitoring systems, and applying performance data; and the use of evidence and learning, including evidence-based design and the uptake of lessons learned." This report assesses progress (or lack thereof) on all these issues.

Figure 1: Situating the evaluation function within UNICEF's broader enabling environment

Source: Created by the peer review consultants based on findings from the peer review

UNICEF has committed more resources and adapted its funding modalities since 2018

In the three years following launch of the 2018 *Evaluation Policy*, UNICEF expenditure on evaluation was reported to have increased from US\$50 million to US\$65 million. Over this same time, UNICEF's programme actuals were also reported to have increased from US\$6.2 billion to US\$7.2 billion. We will come back to an issue later in the report on the need for management to ensure sufficient and predictable funding for the evaluation function. This period also saw the development of the UNICEF Evaluation Pooled Funds, which are said to be "vital to ensure the gains in coverage made at the decentralized level are to be sustainable."¹¹ Pooled Funds are "set up either at the country or global level, come with a contribution of multiple donors co-mingled, earmarked or lightly earmarked to finance joint UN interventions."¹² Additional funding allowed regions to recruit multi-country evaluation specialists and all regions passed the 0.5 per cent funding level by 2021.¹³ However, issues around some lack of clarity regarding governance of the pooled funds as well as criteria for their allocation and use should be clarified.

There has been progress towards the goals of the 2018 Evaluation Policy, though some areas could be further strengthened

The 2018 *Evaluation Policy* intends to set out "the purpose and use of evaluation in UNICEF, provides definitions, norms and standards and outlines governance arrangements and accountabilities as well

¹¹ UNICEF Executive Board – Informal briefing – 24 May 2022 Item 7: Annual report for 2021 on the evaluation function in UNICEF.

¹² <https://www.unicef.org/partnerships/funding/UN-inter-organizational-arrangements>

¹³ Insight Evaluation Function Performance Dashboard as reported in UNICEF Executive Board – 2022 annual session (14-17 June 2022) Item 7: Annual report for 2021 on the evaluation function in UNICEF Reference document: E/ICEF/2022/17 https://www.unicef.org/executiveboard/media/11696/file/2022_AS-Item_7-Evaluation_function-R.McCouch-Presentation-EN-2022.06.06.pdf

as performance standards for the evaluation function. It guides UNICEF staff and partners regarding the organization's requirements for the conduct and use of evaluations. The policy applies to all levels and parts of the organization." The policy is designed to reflect international norms and standards in evaluation and, in particular, the norms and standards defined by UNEG. These norms and standards establish the centrality of utility, credibility and independence. Our documentary review confirms that these norms and standards are explicitly included in an expansive range of formal documentation.

The 2018 *Evaluation Policy* also expects that "evaluations should be properly planned, managed, effectively conducted and quality assured, and the quality of evaluation reports assessed systematically" and, in keeping with the decentralized approach, it is well documented that evaluations should be managed at three levels: the plan for global evaluations, prepared by the Evaluation Office and approved by the Executive Board in support of the UNICEF Strategic Plan and other corporate policies and strategies, which builds upon regional and multi-country evaluations and country-level evaluations and identifies the criteria for the selection and prioritization of corporate evaluations; regional evaluation plans, prepared by regional offices for evaluating multi-country initiatives, which are costed and build upon country-level evaluations and respond to the evaluation needs of the countries in the region; and country Costed Evaluation Plans (CEPs), prepared by country offices with the participation of national governments, which are approved by the Executive Board concurrently with country programme documents (CPDs).¹⁴

An organizational evaluation function is expected to be independent but not isolated; it is expected to support Executive Board, organizational and partner accountability, decision-making, programming and learning. The theory of change describing the intention for how the 2018 *Evaluation Policy* would achieve its intended proximal outcomes and more distal impacts is outlined in Annex 9 of this report. While considering the extent to which the theory of change remains relevant today, we also want to understand whether and how the theory of change has actually been used in practice to support the evaluation function and whether it has been updated. We are therefore interested in the wider set of questions concerning whether the *Evaluation Policy* has contributed to strengthening the evaluation function's capacity to deliver its mandate and UNICEF's organizational capacity to effectively use and benefit from the products of the evaluation function.

The Annual report for 2021 on the evaluation function in UNICEF, published in June 2022, noted that "[d]espite numerous significant challenges, the evaluation function adapted and remained strong." It reported increased expenditure and progress towards meeting the goals of the 2018 *Evaluation Policy*. It reported more submissions, better geographical coverage and more diversified evaluation products. Evaluation quality, as reported through GEROS, remained high in 2021. At the same time, the 2021 Annual Report identified possible areas for improvement. These especially relate to the utility of evaluations, including better use of executive summaries and lessons learned sections. These aspects of good progress, along with a range of areas for improvement, are explored throughout our findings.

The evaluation challenges faced by UNICEF are not unique among organizations

In order to explore whether the challenges facing UNICEF were in some sense specific to the organization, we conducted a light-touch review of the evaluation policies of the World Food Programme, USAID, UNDP and the Asian Development Bank, set out in Annex 10 of this report. We

¹⁴ <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N18/106/37/PDF/N1810637.pdf?OpenElement>, p. 8

found that the challenges facing each agency were very similar but that there were some responses that may have relevance to UNICEF.

In both UN agencies, UNEG norms and standards directly shaped policy and practice. Both UN agencies have an evaluation policy with broad aims that are similar to those of UNICEF (to achieve independent, credible and useful evaluations embedded into the policy and programme cycle, with all evaluations managed in accordance with the United Nation Evaluation Group's Norms & Standards, and an Evaluation Quality Assurance System). All four comparator agencies have governance arrangements that involve an executive board (or equivalent) and a director of evaluation. Each attempts to have some central direction using policy, prioritization, guidelines and leadership but also to ensure some decentralization to encourage ownership and a bottom-up approach to identifying what needs to be evaluated. ADB reports that there are ambiguities in roles and responsibilities that blur the boundary between independence and engagement. WFP, UNDP and USAID all report high levels of use of evaluations, with ADB reporting more weaknesses in this respect. One study in UNDP highlighted the risk that decentralized evaluations could result in low-quality evaluations.¹⁵ Credibility is said to be strengthened by stakeholder engagement as well as technical proficiency and impartiality. Recruitment and retention of quality staff is an issue for all, as was identifying skilled external evaluation consultants.

We identified several interesting practices. We have listed these at the end of Annex 10 for consideration of the UNICEF Evaluation Office.

¹⁵ UNDP (2019). Independent Review of UNDP Evaluation Policy.

FINDINGS

Introduction

Since the first peer review in 2006, UNICEF’s evaluation function has developed greater clarity about what evaluation means, what standards and principles should be applied and what quality standards are acceptable, and has strengthened the professionalization of human resources for evaluation. In 2018, it was stated that the rationale for a revised policy was that the “landscape in which evaluations are conducted has been transformed by wide-ranging global accords, including those on sustainable development (2030 Agenda); disaster risk reduction (Sendai Framework for Disaster Risk Reduction 2015-2030); climate change (Paris Agreement under the United Nations Framework Convention on Climate Change); and financing for development (Addis Ababa Action Agenda of the Third International Conference on Financing for Development).”

In response to these changed circumstances (among other things), the 2017 peer review emphasized the importance of strengthening national evaluation capacities and United Nations joint and system-wide evaluations, “recognizing the benefits of promoting learning within the United Nations system, shared accountability and reduced transaction costs.” The *2018 Evaluation Policy* approved in June 2018 committed UNICEF to significant further improvements, including “provisions related to the geographic coverage of evaluations, quality assurance standards, timeliness and use of evaluations, human and financial resource requirements and the role of managers.” The policy had an associated ‘Procedure on the implementation of the 2018 UNICEF evaluation policy’ that outlines the key responsibilities for the Evaluation Office, Headquarters divisions and offices, regional offices and country offices. The alignment of the policy was described as follows: “evaluation policy is aligned with the Convention on the Rights of the Child and the UNICEF mission statement, with the Charter of the United Nations and humanitarian principles, supports the achievement of the 2030 Agenda for Sustainable Development, the UNEG and international good practices in evaluation.”

The findings of this peer review are organized in the following sections according to the criteria identified by UNEG, which are the six main foci of the review:

- Independence
- Credibility
- Usefulness
- Roles and responsibilities for evaluation
- Use of and follow-up on evaluations
- Overall enabling environment for the evaluation function

Independence

Key findings:

- The principle of maintaining the independence of the evaluation function has high visibility in evaluation guidance and policy documents. The principle of evaluation function independence enjoys strong support among stakeholders. The UNICEF Executive Board, like the Executive Boards of other New York-based UN entities, is placing increased attention on ensuring the independence of the independent oversight functions.
- Particularly at the global level, the independence of the function has been enhanced since 2018, including by appointing senior staff with extensive evaluation expertise, and by introducing a specific maximum term for the Director of Evaluation’s service. Independence at

the global level should be sustained through regular direct reporting of the Director of Evaluation to the Executive Director and direct access to the Executive Board.

- Key risks to the independence of the evaluation function, particularly at the decentralized level, were identified. These include the formal structure of roles and accountabilities and evaluation funding modalities, which neither enforce nor overwhelmingly incentivize behavioural independence. This results in variation in the level of actual independence in decentralized settings.
- While the independence of the evaluation function is rightly valued as among the key enablers of credibility and utility and as crucial to oversight and accountability, there will be times, as the operating environment of UNICEF evolves and becomes more complex, when pursuing some more embedded and collaborative approaches to the design and conduct of evaluations may contribute towards strengthening the relevance and utility of findings and recommendations. Such approaches, however, should not detract from the real or perceived ability of the evaluation function to maintain and exercise its independence.

Definition of independence

Independence describes how well placed the evaluation function and its personnel are to set their priorities and deliver their workplans (including use of financial and engagement of human resources) free from interference or improper influence. According to the 2018 Evaluation Policy : “Management arrangements for each evaluation should ensure independence and impartiality. The designated evaluation manager should not be part of the team that designs and/or manages the implementation of the assessed policy, plan, or programme. The designated evaluation manager supervises the selection of the evaluation team, manages the consultants” and “[k]ey stakeholders, including excluded groups and, as appropriate, children and young people, should be engaged at relevant points, starting with the design phase. The involvement of children and young people should follow appropriate ethical guidelines and requires approval in accordance with UNICEF procedure. The evaluation team should be selected through an open, transparent process, with balance in terms of geographical and gender diversity, and should include professionals from the region or country concerned in the evaluation.”

The 2017 peer review found that “behavioural and organizational independence is short of being satisfactory, both for the Evaluation Office and for the decentralized evaluation function. 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.”

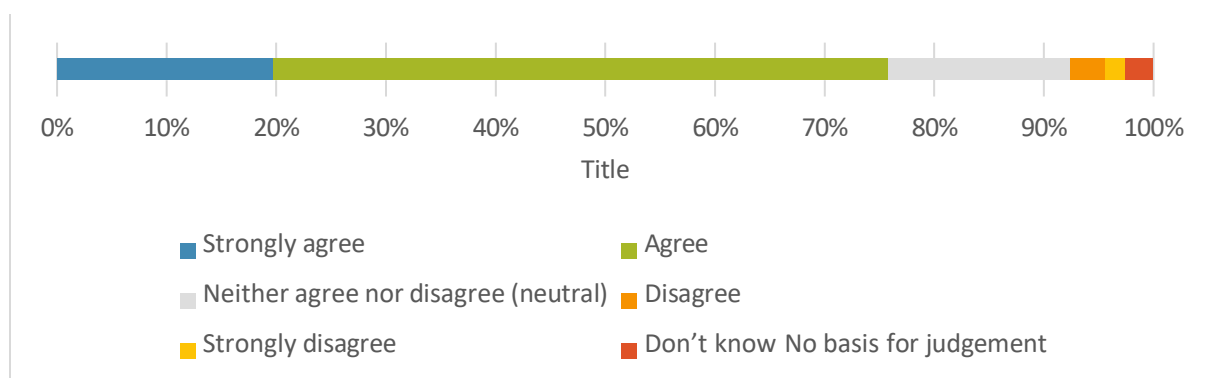
The principle of independent evaluation enjoys wide support within UNICEF

This peer review has found that the principle of maintaining the independence of the evaluation function has high visibility in evaluation guidance and policy documents. The principle of independence also enjoys strong support among stakeholders and enjoys strong support within UNICEF. Throughout our consultations with UNICEF staff and external stakeholders, explanations for why the principle of independence was so valued highlighted the importance – particularly within a large, hierarchical organization such as UNICEF, where staff rotate frequently and may be more reluctant to provide critical analysis of the organization’s programmes and operations.

Independence should be further optimized by addressing certain structural and behavioural factors

Most UNICEF and external stakeholders agreed that, at the global level, the independence of the function has been significantly enhanced since 2018. Survey data show over three quarters of practitioners agree that current models of commissioning and management support the impartiality of evaluation, while almost the same proportion of senior management agree that evaluations are carried out without undue influence by any party. It is notable and welcome that the independent evaluation consultants interviewed for this peer review felt they had largely been allowed to conduct their work without undue influence or interference.

Figure 2: Practitioners' survey: To what extent do you agree or disagree that the current model of commissioning and managing evaluations through external parties supports impartiality of the evaluation process, without the undue influence of end users



At headquarters level, we welcome that evaluation roles are clearly demarcated from other functions. We also welcome that the *2018 Evaluation Policy* requires the Director of Evaluation to have specific evaluation expertise and asserts avoidance of conflict of interest by requiring the role to have a specific maximum term, after which the post-holder is barred from further employment in the organization. The commitment to strengthening the independence of the function by appointing senior staff, including the Director of Evaluation, with substantial evaluation expertise has now been achieved. It is essential that corporate mechanisms ensure that the conflict-of-interest policies are upheld to help to ensure that there are no real or perceived impediments to the independence in respect of the evaluation function vis-à-vis the role of the Director of Evaluation. Independence at the global level should be sustained through regular direct reporting of the Director of Evaluation to the Executive Director and through direct access to the Executive Board.

Key risks to the independence of the evaluation function were identified, however, at the decentralized level. These include the formal structure of roles and accountabilities and funding modalities for evaluation, which neither enforce nor overwhelmingly incentivize behavioural independence. This results in variation in the level of independence in decentralized settings.

At the regional level, the roles of Regional Evaluation Adviser (REA) were created as a result of the *2018 Evaluation Policy*, and we welcome that appointments to REA posts are highly credible in terms of their level of evaluation expertise. One risk, however, is that the REA role currently has a formal reporting line to regional Directors and a 'dotted line' to the Evaluation Director. Overall, this organizational structure does not guarantee that REAs are consistently protected from potential pressures from regional structures that may constrain their independence. As a result of this structure, the independence of REAs could potentially be affected by behavioural factors, including

the priorities of Regional Directors. A diverse range of sources indicated that such risks to independence were heightened in some regions, whereas, in others, no risk materialized. As we explore in further detail in the section below on 'Roles and responsibilities', ensuring a consistent level of independence at the regional level may require more a clearly demarcated structure of roles and accountability.

At the country level, the peer review heard from a range of sources that the independence of the evaluation function is also subject to inconsistency. It was widely reported that, in practice, the extent of the independence of evaluation differs between country offices and is (similarly to regional levels) linked strongly to the perspectives and priorities of individuals in leadership roles. Compared with HQ, understanding of roles and accountabilities in relation to independence is also variable. In addition to the lack of clearly demarcated accountabilities, however, another key issue cited by country office staff as affecting evaluation independence was evaluation budgeting modalities. The central problem at the country office level – affecting efficiency and effectiveness as well as independence – is that budgets for evaluations are not predictable and may not always be set aside. We were told that funding is released at intervals and is subject to the variations of the organization's wider resourcing and budgeting. This means country office evaluation staff – and also REAs – face difficulties in both planning and implementing evaluations, as they await security of funding. Some country offices evaluation professionals that we consulted viewed this as their greatest challenge and one that consumed much of their time.

Particularly at the decentralized level, then, both the structure of roles and accountabilities and funding modalities appear as potential risks to the optimal independence of the evaluation function at regional and country levels and in the context of competition for programming resources. As the structure of roles and accountabilities at decentralized levels does not enforce behaviours that support independence (but, rather, the independence of the evaluation function at these levels may be linked to the priorities of regional and country level leaders), the decentralized evaluation function thus appears somewhat vulnerable to potential pressures and preferences.

While much of the scaffolding, in terms of policy and guidance, to support the *principle* of independence is in place, in the section below on roles and responsibilities, we highlight ways in which *actual structural and behavioural* independence may be strengthened. Our discussion of evaluation funding (in the section below on the enabling environment for evaluation) also suggests ways in which alternative modalities of evaluation funding may better support independence.

[Pursuing some more embedded and more collaborative, synergistic approaches, where relevant and appropriate, could help to improve relevance and utility](#)

While independence of the evaluation function is an essential cornerstone that must be protected, it is important to state that more embedded and collaborative, synergistic approaches to evaluation, where relevant and appropriate, also offer value – and can also uphold the principle of independence if well managed. Conversely, and in general, there is a risk that independent evaluations that are not managed to gain sufficient buy-in from management may be viewed as irrelevant and struggle to promote ownership and championing of findings and recommendations. A widely recognized norm within the global evaluation community is that the independence of evaluation should not mean its isolation. Many stakeholders who were consulted highlighted that the UNICEF evaluation function should be independent but that it could be accompanied by further engagement with stakeholders to ensure relevance and usefulness. The goal of harnessing UNICEF's evaluation function to meet high-level strategic priorities should involve collaboration among key knowledge functions of the organization and external agencies to maximize potential synergies.

Responsibility for achieving this cannot rest with the evaluation function alone: it requires and depends on mutual willingness to collaborate among potential partners and evaluation users. Across all levels of UNICEF, there is an internal market for people’s attention. As we explore in further detail in the section below on the enabling environment (and particularly the subsections on strengthening UNICEF’s evaluation culture and maximizing the benefits of evaluation through collaboration with other functions), there is potential to continue optimizing the relevance and value of evaluation within UNICEF. The success of endeavours to achieve this on the part of the evaluation office will depend, however, on aspects of the enabling environment that are in part outside its sphere of full control.

Credibility

Key findings:

- GEROS data suggest a strong and improving situation in relation to the quality and geographical coverage of UNICEF evaluations, as well as incremental improvements to the integration of disability and gender equality considerations with evaluations. GEROS itself has emerged as a key enabler of quality and credibility.
- Most senior managers and evaluation practitioners in UNICEF regard evaluation as credible, though a sizable majority suggest there is room for improvement.
- The credibility of the evaluation function has been strengthened by investments in human and financial resources.
- Progress on credibility is affected by a range of factors, some of which are within the control of the evaluation function and some of which require UNICEF management action. Factors of the first category include: improving the consistency in the quality of decentralized evaluations and their methodological approaches and better focusing the portfolio of evaluations on UNICEF’s overarching strategic priorities and on areas where evidence is lacking or most needed. UNICEF management action is required to ensure both the availability of high-quality monitoring data and effective funding modalities that support evaluation planning and implementation as well as application of learning within workplans.

Definition of credibility

Evaluations should be credible, meaning that evaluation products, processes, analyses and formulations of conclusions and recommendations should demonstrate impartiality and appropriate engagement with duty bearers and rights holders and should be based on sound methods and analysis. In addition, the individuals involved in evaluation should carry personal credibility both in relation to their technical skills and to their skills such as stakeholder engagement and awareness of the organizational and disciplinary context of evaluations.

According to the 2017 Peer Review Panel, “[T]he Peer review’s assessment of credibility was short of being satisfactory, both for the Evaluation Office and for the decentralized evaluation function. This was 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.” These factors have improved with the implementation of the revised policy.

There have been measurable improvements in evaluation quality

The 2018 requires quality assessment of all evaluation reports, stating, “The quality of all evaluation reports is assessed by a specialist external to UNICEF. The Evaluation Office is responsible for ensuring that a credible quality-assessment process is established and that its results are shared with the commissioning office and made available to UNICEF management and the Executive Board. Periodically, the Office will undertake reviews of the quality-assessment system for all levels of the organization.” Surveys of senior management and evaluation practitioners suggest a fairly mixed view of the credibility of evaluations, though the majority do hold a positive view, and GEROS data demonstrate measurable improvements in quality and geographical coverage.

Survey results show that, while senior management (across a variety of offices and divisions) and evaluation practitioners do not universally endorse the credibility of evaluation, the majority agree that evaluations are largely credible, with room for improvement. The senior management survey found that 66 per cent agreed that evaluations were well designed and 57 per cent agreed that analysis of evidence was accurate and robust. It is notable, however, that only 48 per cent agreed that evaluations were perceived as authoritative by the partners they worked with. Almost three quarters of practitioners agreed that current commissioning and management models support credibility and rigour.

Figure 3: Practitioners’ survey: To what extent do you agree or disagree that the current model of commissioning and managing evaluations through external parties supports credibility and rigour of analysis?

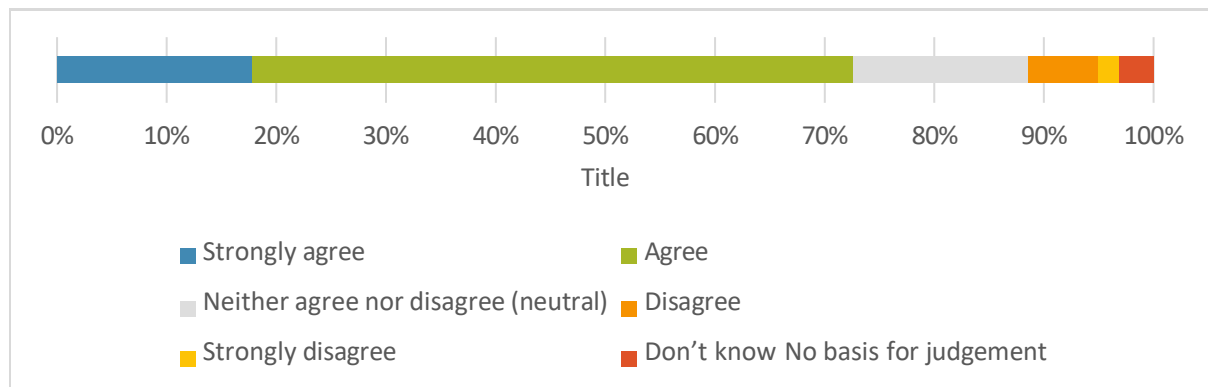
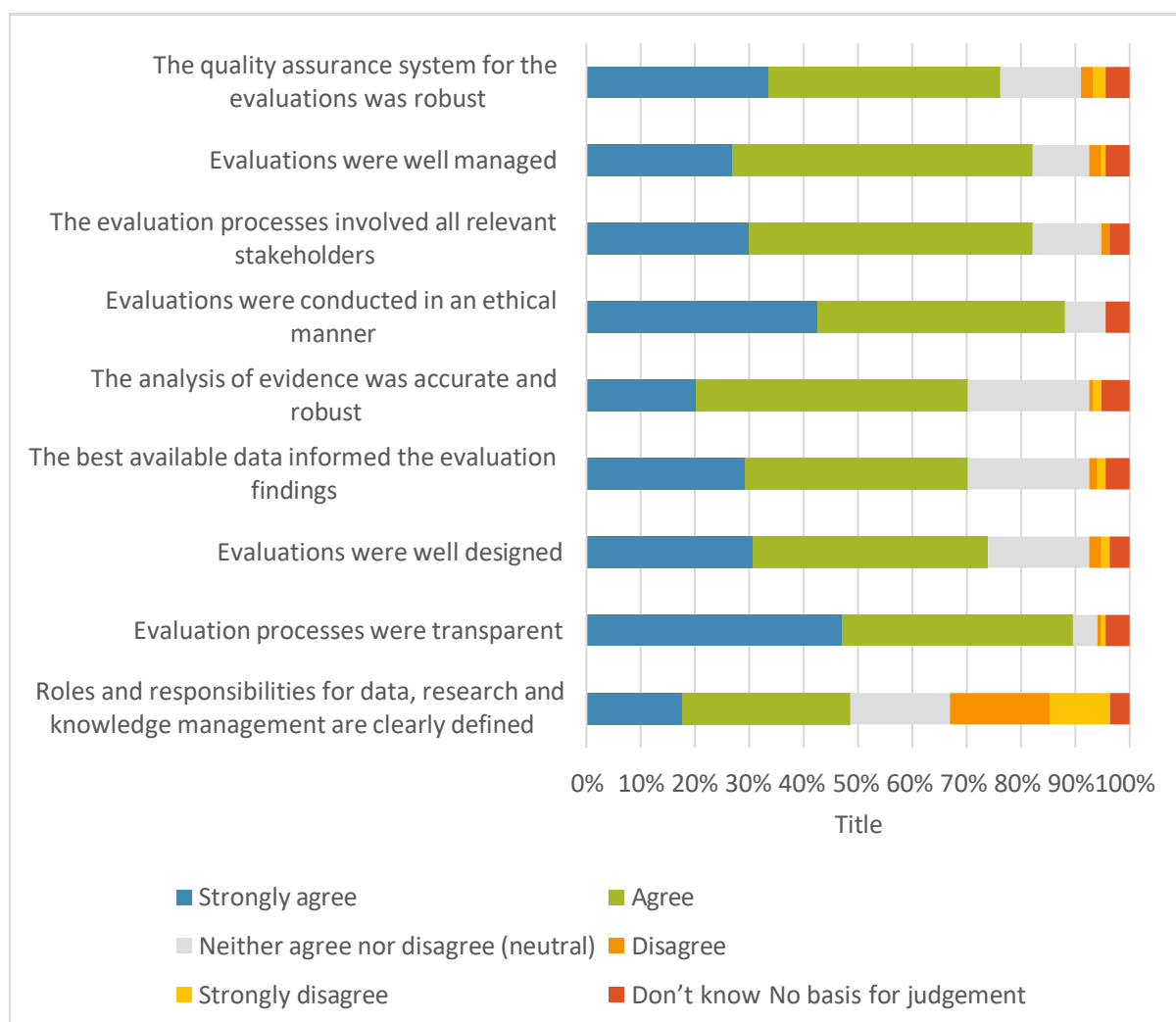


Figure 4: Practitioners' survey: To what extent do you agree or disagree with the following statements about the credibility of the evaluations conducted in your office/division?



The 2020 ‘Study on Influential Evaluations in UNICEF’, commissioned by the UNICEF Evaluation Office, also highlights that the organizational and policy setting facilitates, at its best, highly credible and influential evaluations with both internal and external influence. (Generally, the study provides a careful reflection of current issues in evaluation and their implications for UNICEF and stands as an example of good practice by the UNICEF evaluation function. It may be useful for other large evaluation functions to conduct similar studies within their own organizations.) The Study suggests that influence and credibility are most often associated with strong contextual understanding, evaluation design, stakeholder engagement, and trust and collaboration.

GEROS¹⁶ data suggest a strong and improving situation

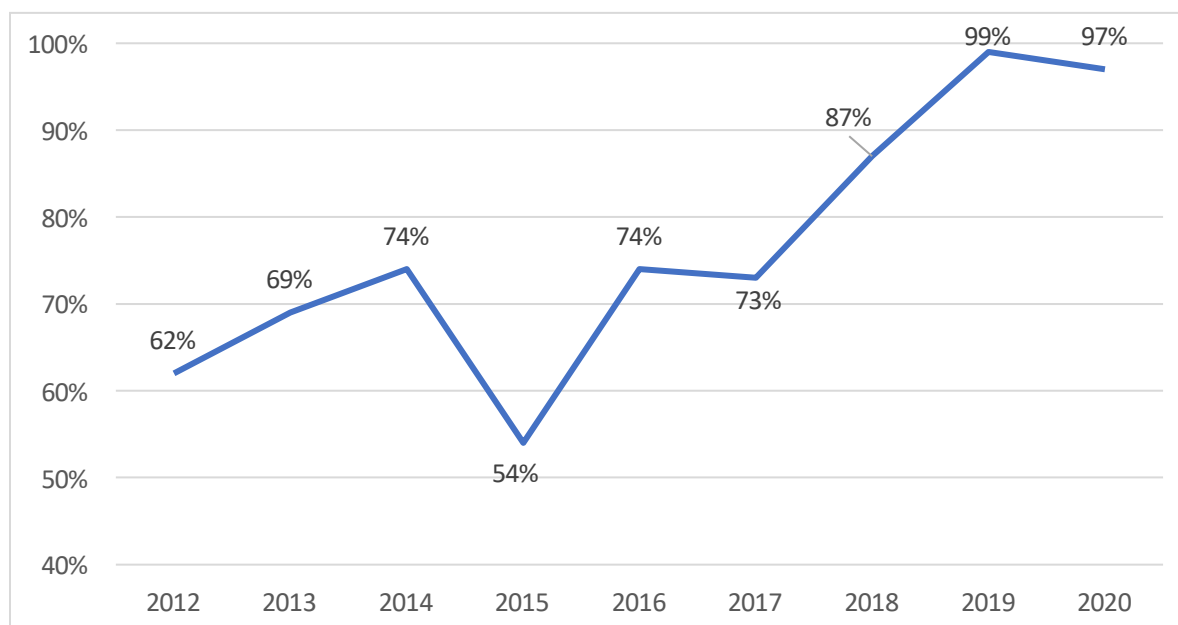
As credibility is underpinned by quality, data produced through GEROS importantly suggest a strong and improving situation in relation to the quality and geographical coverage of UNICEF evaluations. The use of GEROS was also, itself, viewed by stakeholders consulted for this peer review as an

¹⁶ GEROS was established in 2010 to provide a holistic system to ensure the quality and credibility of evaluation reports across the whole organization, to meet UNEG standards and to support institutional learning and the use of evidence.

established enabler of credibility. Based on the adage ‘what gets measured gets done’, this suggests that repeated measurement of quality has focused attention on improving it.¹⁷

As shown in the diagram below, GEROS scoring improved significantly across the 2010s and, following a dip in 2015, eventually increased to 97 per cent of evaluations being rated ‘good’ (either satisfactory, highly satisfactory or exceptional) in 2020.

Figure 5: Percentage of evaluations with GEROS ratings of adequate or better (2012-2020)



The peer review consultants examined a sample of reports that were rated as exceptional according to the UNICEF GEROS system.¹⁸ Most of these were summative, formative or both, and used mixed-methods. The GEROS reviewers praised these reports for detailed and considered methodology and for data collection that drew on a wide range of stakeholders. The reports offered both high-level and specific practical recommendations that took account of the practical contexts of programme delivery. Almost all reports rated exceptional were praised specifically for their integration and consideration of gender analysis. This GEROS evidence (along with the findings of the 2020 *Study on Influential Evaluations in UNICEF*) supports the conclusion that the current evaluation function can deliver very high-quality and highly credible evaluations.

GEROS reports were not available for reports rated as unsatisfactory, but we were able to examine a sample of reports rated as satisfactory. GEROS Reviewers praised these reports for detailed and comprehensive analysis. However, some were also criticized for being poorly formatted and difficult

¹⁷ This is as envisaged by UNICEF <https://www.unicef.org/evaluation/global-evaluation-reports-oversight-system-geros>

¹⁸ The GE OS Quality rating scale ranges from ‘Exceptional’ (‘Exemplifies UNICEF UNEG standards for evaluation report’ and ‘Decision makers may use the evaluation with a high degree of confidence’) to ‘Unsatisfactory’ (‘Does not sufficiently meet the UNICEF UNEG standards for evaluation reports’ and ‘Decision makers cannot rely on the evaluation’). Intermediate points are ‘highly satisfactory’, ‘satisfactory’ and ‘fair’. It should be noted that the bar of ‘Exceptional’ is not set at a high level. However, it should also be recognized that the methodology for arriving at judgements is robust and extensive. See: <https://www.alnap.org/system/files/content/resource/files/main/geros-unicef.pdf>

to follow or for offering vague, high-level recommendations instead of actionable recommendations targeted at specific stakeholders.

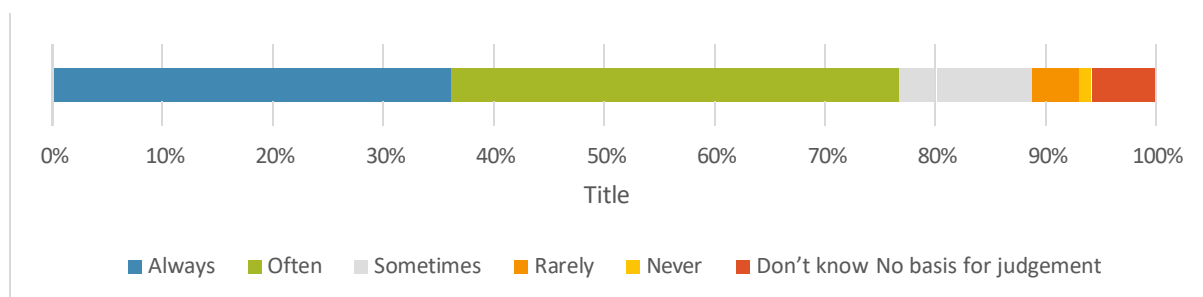
In addition to improvements in GEROS quality ratings, there is also evidence that the evaluation function is producing more evaluation products, more real-time assessments and more evaluations integrating disability and gender equality in the scope of analysis.

Of 178 evaluative products submitted in 2021 and included in GEROS data, 14 were conducted at headquarters level and 164 were conducted at the decentralized level; this represents 23 more evaluative products than in 2020.¹⁹ In addition, six evaluability assessments and two reviews were conducted that aimed to provide timely feedback on implementation of education management information and monitoring systems. Fourteen real-time assessments were submitted in 2021, compared with five in 2020; all of these were related to the COVID-19 pandemic and contributed to the evidence base to inform country and regional COVID-19 responses. In response to COVID-19, there was also increased use of remote techniques and other workarounds.

At the same time, the number of evaluations covering disability in evaluations continued to increase as offices mainstreamed disability across all policies and programmes. Importantly, the Evaluation Office undertook a meta-synthesis in September 2022 of the inclusion of disability in evaluations, finding 56 good-quality disability inclusive evaluations among the 2018-2021 sample.²⁰ These results were cross-checked with the (self-reported) disability tag used in the Evidence Information Systems Integration database (EISI), which uncovered inconsistencies and suggested limitations in understanding of what disability-inclusive evaluations are. This was felt by the Evaluation Office to support the need to develop further policy guidance, which it then produced.

GEROS shows consistent improvement in the integration of gender equality and the empowerment of women in the scope of analysis of evaluations, while the evaluation practitioner survey suggests the view that over three quarters of recent evaluations include an assessment of issues related to gender sensitivity or equality. Less progress was indicated in GEROS on the use of gender-responsive methodology, methods and tools, and data analysis and in the reflection of gender analysis in the evaluation findings, conclusions and recommendations.

Figure 6: Practitioners' survey: How frequently did evaluations include an assessment of issues related to gender sensitivity or equality in the past 24 months?



The overall picture from GEROS data is of an evaluation function that is continuing to progress towards a credible balance, and improved quality, of evaluations that provide managers with access to a range of tools and evidence to inform programming. These improvements are also apparent in

¹⁹ [Evaluation reports | UNICEF Evaluation in UNICEF](#)

²⁰ <https://www.unicef.org/evaluation/media/2866/file/Disability-Inclusive%20Evaluations%20in%20UNICEF:%20Guideline%20for%20Achieving%20UNDIS%20Standards.pdf>

the Plan for global evaluations, 2022-2025, which shows a welcome willingness to learn from previous reviews.

Financial and human resources strengthen the credibility of the evaluation function

In terms of the factors supporting the improvements discussed above, the commitment to spending globally 1 per cent of UNICEF's programme resources on evaluation focuses attention and provides a lever for encouraging resources to be allocated to evaluation. While this target has not yet been achieved, from 2018 to 2021, UNICEF reported that it had recorded an additional US\$15 million spent on the evaluation function.²¹ In the three years following launch of the 2018 *Evaluation Policy*, UNICEF expenditure on evaluation reportedly increased from US\$50 million to US\$65 million. It should be noted, however, that, over this same time, UNICEF's programme actuals were also reported to have increased from US\$6.2 billion to US\$7.2 billion.

In addition to increased financial resources, this peer review learned that investment in human resources and the level of in-house evaluation expertise, particularly in relation to Evaluation Office staff and Regional Evaluation Advisers, have supported improvements in credibility.

Further progress on credibility is limited by a range of factors

While there has been welcome progress on credibility since 2018, this review also found some areas for improvement. Some areas are more within the Evaluation Office's control and some require management action. Factors of the first category include: improving the consistency in the quality of decentralized evaluations and their methodological approaches; and better focusing the portfolio of evaluations on UNICEF's overarching strategic priorities and on areas where evidence is lacking or most needed. UNICEF management action is required to ensure the availability of high-quality monitoring data and to ensure effective funding modalities that support evaluation planning and implementation as well as application of learning within workplans.

Improving consistency in decentralized evaluation quality and the appropriateness of methodological approaches

While GEROS quality assessments show a strong and improving situation with regard to evaluation quality and while some stakeholders consulted for this peer review praised the quality of particular UNICEF evaluations, the overall view from stakeholders was that the quality of evaluations and the appropriateness of methodological approaches at regional and country levels were not always consistent.

Particularly among some UNICEF stakeholders at the frontline of commissioning and managing decentralized evaluations, concerns were expressed about the variable quality of external evaluation consultants, often because budgets were apparently too low to attract highly skilled and knowledgeable consultants. In some cases, evaluations were less than satisfactory due to overly expansive terms of reference (with, for example, too many evaluation questions to allow each question to be adequately answered with the resources available). Other problems discussed by staff at decentralized levels included how to ensure that evaluators engaged by UNICEF understand UNICEF systems and the UNICEF approach to results, human rights-based approaches and gender perspectives. This 'frontline' view was echoed by users of evaluation in Headquarters, who suggested there was a wide range of quality in products and external providers, a heavier emphasis on quantity over quality and too many small-scale evaluations.

²¹ https://www.unicef.org/executiveboard/media/11696/file/2022_AS-Item_7-Evaluation_function-R.McCouch-Presentation-EN-2022.06.06.pdf

Some stakeholders viewed part of the solution to these issues as involving building on the experience of providers with Long-Term Agreements – though, as a caution on this approach, going back to the same evaluators again and again may also contribute to somewhat of an ‘echo chamber’, repetitive findings and recommendations, and lack of methodological innovation.

Another solution, which could be pursued alongside this, would be to build in-house evaluation and evaluation management capability, so that evaluation staff in UNICEF are better placed to manage, lead and conduct evaluations themselves (with support and input from external experts where needed). This has already been piloted and Evaluation Office staff report many lessons were learned. An Evaluation Office staff member led a real-time COVID evaluation exercise and Evaluation Office staff have been embedded as team members in at least five evaluations since 2018. This could be built upon by further developing relevant evaluation skills and human resources at the country office level. Staff being more involved in conducting evaluations provides potential advantages but also comes with a time commitment that may necessitate offsetting other activities.

The degree of evaluation expertise within country offices was usually said by consulted stakeholders to vary from country to country. Continued investment in the development of evaluation knowledge and skills at this level would be welcome, particularly to enable better design and management of country-level evaluations and so to improve the credibility of decentralized evaluation. The practitioners’ survey found that the top five factors viewed by practitioners as affecting the performance of staff in the evaluation function and the quantity and quality of evaluations included both access to high-quality evaluation consultants and technical expertise in evaluation.

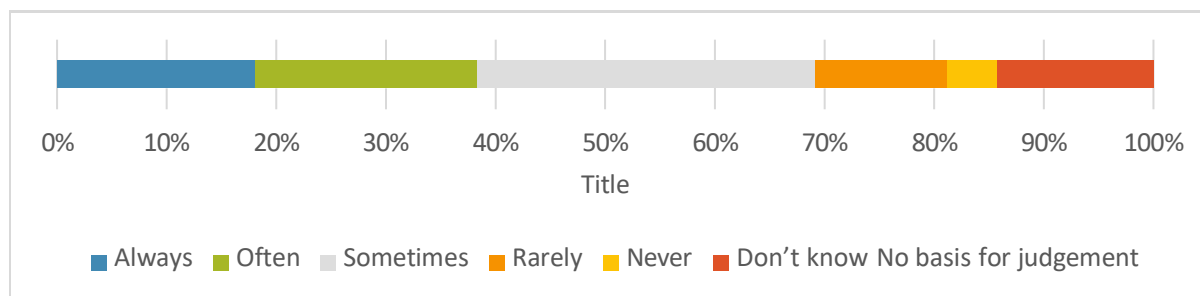
In relation to evaluation methodologies, future plans outlined to the Executive Board by the UNICEF Evaluation Director in 2022 include focusing more on outcome and impact evaluations, and this is also reflected in the outcome-focused Strategic Plan and Evaluation of Impact Strategy. Many stakeholders were very positive about the prospect of increasing focus on impact evaluation, which can be viewed as a key step along the route to gaining more robust understanding of what works to improve outcomes for children. Impact evaluations are focused on the short and long terms and on the positive and negative consequences of an intervention and, as such, can be very valuable.²²

Some cautions were expressed around the limitations of impact evaluation and the need to manage expectations around the conditions under which they are feasible and appropriate. Impact evaluation often requires a high level of resourcing, involves long timescales and may not be appropriate in contexts where there are low levels of programme fidelity, where complex and changing circumstances make attribution difficult (because programme effects are hard to disentangle from confounding factors), where establishing the counter-factual is technically challenging or practically impossible or where outcomes do not lend themselves to comparable quantitative measures or are not expected within the time period that can be evaluated. While improving the availability of high-quality monitoring data would make impact evaluations less expensive and easier to deliver, impact evaluations also tend to be relatively expensive and resource-intensive. The use of impact feasibility assessments may help the evaluation function to identify where impact evaluation may be less appropriate and useful than other methods for assessing impact, such as contribution analysis or process tracing.

²² See: <https://www.betterevaluation.org/methods-approaches/themes/impact-evaluation#:~:text=An%20impact%20evaluation%20should%20only,about%20the%20intervention%20under%20investigation.>

Future plans outlined to the Executive Board by the UNICEF Evaluation Director include expanding the range of evaluative exercises to inform the work of the organization. In line with this, stakeholders consulted for this peer review noted that, alongside and without detracting from the emphasis on impact evaluation, UNICEF should retain and strengthen its focus on participatory approaches that involve children and young people not only as sources of data but also as key contributors to evaluation planning, design, analysis and production of findings and recommendations, where appropriate and with the necessary ethical standards implemented. Given UNICEF's mandate and the 2022-25 Strategic Plan, youth participation also has an important role to play in the organization's work, including evaluation. While UNICEF has resources that provide guidance on participatory and co-production approaches, efforts to implement these at the global, regional and country levels could be strengthened. It is important to recognize that such efforts would need to be adequately funded, given that it takes investment of time and expertise for children and young people to participate meaningfully and safely. Given UNICEF's significant global reach and influence, this could be part of a wider move to position itself as a leading actor in the conduct and promotion of child-sensitive inclusive and participatory approaches to evaluation.

Figure 7: Practitioners' survey: How frequently was data collected about the perspectives of children in your office/division in the past 24 months?



Harnessing the evaluation portfolio to strategic priorities and to what matters the most
 One concern for credibility concerns what is being evaluated, i.e., the coverage and treatment of key issues within the portfolio of evaluations. The *2020 Study on Influential Evaluations in UNICEF* comments on the balance of the evaluation portfolio, noting (for example) the low level of strategy-level evaluations and systems evaluations, the relatively few evaluations of pilots and innovations, and the limited attention given to global drivers such as climate change and synthesizing evaluations. The *Plan for Global Evaluations, 2022-2025*, building on the UNICEF Strategic Plan, identifies how UNICEF plans to address many of these issues in its proposed distribution of evaluation topics at all levels.

Several internal and external stakeholders expressed the view that UNICEF could improve the strategic focus of its evaluation portfolio, steering evaluation to where it adds most value, to improve credibility and relevance. Adding value in this context means providing new evidence where existing research is absent or weak and where new such knowledge would meet operational needs and help deliver organizational priorities. Existing evaluation selection processes at country and regional levels help achieve geographical and thematic balance but are not always driven by a rationale to create this added value.

At the same time, the evaluation portfolio also must reflect the interests of donors and national partners. It was not known how much of the decentralized evaluation portfolio takes place because of donor funding compliance-related provisions and how much is driven by an active interest (or need) to know for accountability and learning. Across the decentralized function, the proportion of

donor-driven evaluations is non-negligible, but other factors, determined at a decentralized level, are also relevant.

Evaluation teams commission both demand-driven evaluations (i.e., the areas stakeholders *want* to evaluate) and evidence-, risk- and opportunity-based evaluations (i.e., those areas that more systematic, objective and impartial analysis would *point* the team toward evaluating because they are of greatest strategic importance to the organization). Due to the decentralized nature of the evaluation function, decisions about what to evaluate at country level involve a negotiation between several competing interests that include donor requirements, interests and priorities of the UNICEF country representative and the views of the Regional Evaluation Advisors (REA) and Evaluation Office, where they exist.

Within this broad dynamic, some country representatives are more interested than others in evaluation and act as champions for evaluation. But strategic assessment of what should be evaluated is inconsistent at the decentralized level. Qualitative evidence collected for this peer review suggests that country-level evaluation portfolios may, at times, include areas that are not of high strategic priority or not include areas that are of the highest strategic priority.

Donors and other external partners consulted for this peer review noted that, in addition to programme evaluations (which are valued by donors as demonstrating accountability for how funding is spent), they would very much value more work at the corporate level to synthesize findings from across several evaluations on topics relevant to them and to present this in accessible formats (such as 10-page reports).

Improving the availability of monitoring data

The quality of monitoring data was identified as a key challenge affecting the credibility and quality of evaluation, with some cases of limited availability of monitoring data and other cases of limited use of existing monitoring data in evaluation design. For example, there was concern that, where programmes collect limited monitoring data, this places restraints on the methodologies that can be employed (particularly in relation to evaluating impact) and less ambitious methodologies are therefore pursued. Better monitoring data (in addition to tightly controlled interventions) will be needed if UNICEF is to deliver more impact evaluations.

It was also noted that the scope for real-time evaluations and other approaches using existing data was in some cases compromised by the availability of reliable data.²³ Discussions revealed an untapped potential for significantly strengthening the role of digital support for real-time evaluation. This digital function would also strengthen the ease and effectiveness of conducting impact evaluations. However, delivering this successfully would require a strong enabling environment. The plan for global evaluations also suggests a movement towards more integrated working within the evaluation function and with other partners towards a more digitally enabled, holistic collaboration focused on achieving SDGs. This is expected to include greater use of non-traditional evaluations. If this progresses further, it would represent a further opportunity for potentially strengthening evaluation efficiency and effectiveness.

Ensuring that funding modalities support effective evaluation for accountability and learning

While the increase in evaluation funding in recent years is welcome and no stakeholder we consulted suggested that the 1 per cent target was insufficient to deliver credible evaluations, this

²³ We note that there is no agreed definition of 'real-time evaluation' and respondents within UNICEF use the term variously. However, it always involves collecting data more or less in real time and using these data to course correct or to adapt an intervention.

peer review has found that the current funding modalities in practice do not consistently or predictably support effective evaluation for accountability and learning. The current geographical and topical breadth of evaluations, the range of guidance and support, and the investments in human resources suggest that reaching the 1 per cent target may be sufficient to support a credible evaluation function and thereby credible evaluations. It does not appear to be the global sum available for evaluation but rather the funding modalities, and associated roles and responsibilities, that constrain credibility.

We found uncertainties and confusions about whether the 1 per cent target funds were intended to be spread evenly across all programmes or to create a 'global total'. It was also unclear to most of those we consulted what sanctions (if any) might follow if, at country or regional levels, targets for evaluation spending are missed.

It is management's responsibility to ensure that the evaluation function has sufficient and predictable funding to carry out its work. While increased financial resources have helped support evaluation staffing (including multi-country evaluation specialists) and the delivery of evaluations, evaluation staff, particularly those at the country level, report that the lack of certainty about the availability of funding, and the variable priority extended to evaluations, means that staff do not feel able to plan effectively. Evaluation pooled funding was established in 2018 and appears to have been an effective tool to support decentralized evaluations. However, once again, it is seen to be uneven in its effects and there was lack of clarity in terms of its governance, criteria for allocation or use of the funding. We were told that it is released in batches (per quarter) and REAs and MCES are uncertain about how much will be released and that unpredictable funding undermines timeliness and relevance (both of which relate to credibility). Furthermore, the Director of Evaluation has fiduciary duty for the evaluation pooled funding, but we understand that Regional Directors have authority for how it is used. Thus, exploring ways to improve the stability and predictability of funding could help UNICEF staff, particularly at the decentralized levels, to plan and implement programmes of evaluation that the evaluation function deems appropriate.

A further suggestion for funding sustainability was that a formalized commitment to ensuring that a certain percentage of all earmarked funds is allocated to evaluation may help to ensure that evaluation funding keeps up with the scale of programming going forward.

Utility

Key findings:

- The utility of evaluations is a result of a relationship between demand (interest in, awareness of and capacity to use recommendations) and supply (the communication of relevant, actionable and evidence-based findings and recommendations). Not all of this is within the control of the evaluation function.
- The review found important barriers to utility along the evaluation lifecycle from conceptualizing the issues to be evaluated, to recruiting consultants, to communicating results.
- One way that utility could be strengthened would be by including more voices when scoping the issues and identifying recommendations, as well as by ensuring that productions of evaluation products are in easily accessible formats.

- Practitioners expressed concerns about the quality of available external providers willing to work for the rates available, who may not sufficiently understand the needs and ways of working of UNICEF.
- While management responses to evaluation are nearly always produced, this is not in itself proof that evaluations have been used to improve policy, programming, practice and outcomes for children.

Definition of utility

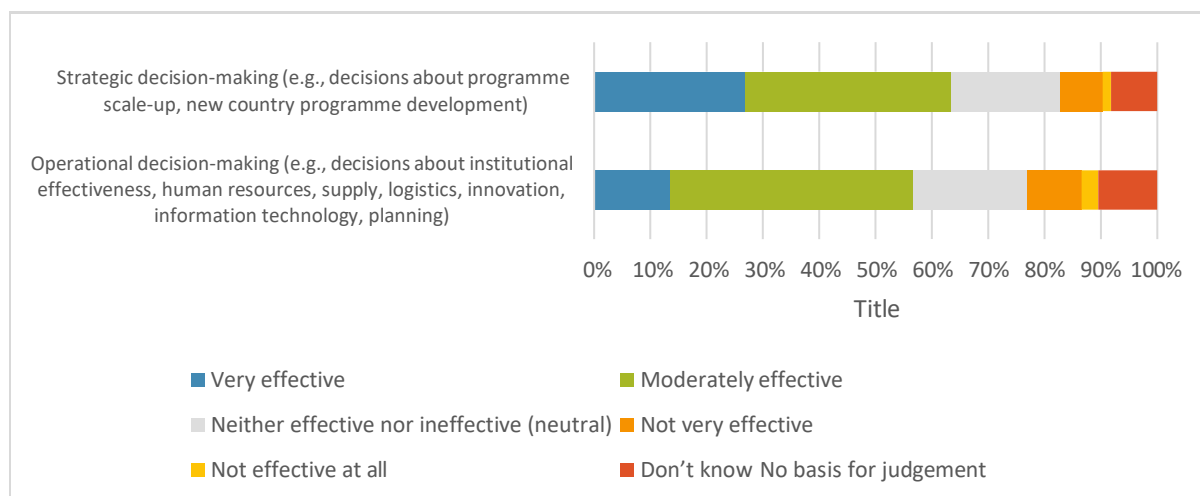
The intention of UNICEF is that every evaluation should support learning or accountability, or both. UNEG standards assert that, at the earliest stages of commissioning an evaluation, the purpose of the evaluation should be clear. As noted in the previous section, to achieve utility, evaluations should be credible and relevant and the knowledge produced should be mobilized to inform decision-making and accountabilities. Utility can therefore be strengthened (or undermined) at any stage along the evaluation lifecycle – from initially conceptualizing the case for an evaluation, to developing the terms of reference, to managing the evaluation, through to mobilizing the results and knowledge produced. Achieving utility begins with identifying evaluation topics where the utility is likely to be greatest. It also depends on ensuring an effective organizational culture of evaluation, cultivated and sustained by management.

The prior peer review found “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.”

Views on the utility of evaluations are mixed

A fairly mixed view of the utility of UNICEF evaluations is reflected in the recent practitioner and senior management surveys, although the majority of stakeholders feel they are engaging with evaluations and find them useful. Among senior management survey respondents, 76 per cent felt evaluative exercises had been useful to their office or division, citing the three main reasons for this as relevance to programmatic needs, validity of findings and recommendations, and validation through consultations with stakeholders. Nonetheless, these results show some room for improvement, with almost a quarter of senior management not reporting that they found evaluation useful or very useful. Where they are not using evaluations, it is most commonly due to the recency of the recommendations or the lack of interest from stakeholders. Notably, less than two thirds of practitioners reported that evaluation products had been effective in meeting operational and strategic decision-making purposes in their office or division.

Figure 8: Practitioners' survey: To what extent have the final evaluation products been effective in meeting the following purposes in your office/division?



Similarly, stakeholders consulted through interviews and focus groups for this peer review also had mixed views on evaluation utility and mixed levels of confidence that evaluations are well used. Some stakeholders suggested that linking the evaluation function more closely to those engaged in programming would enable more lessons for evaluation to be applied. We were also told that strategic oversight of the evaluation portfolio as a whole could be improved to ensure that it supports the priorities of the UNICEF strategic plan. Evaluation evidence should be present and well integrated to be used within strategic and programmatic decision-making processes.

Delivering organizational and wider utility require incremental improvements along the evaluation journey and an effective culture of evaluation

The 2022 evaluation plan provides a greater steer on ensuring the utility of the evaluation function, which is very welcome. Central issues identified by this peer review as requiring ongoing attention include the need to strengthen: the selection of evaluation topics; the planning and management of evaluations; the format and accessibility of evaluation products; and the response to and application of evidence within management decision-making.

Improving selection of evaluation topics

As noted in the 2017 peer review, “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.” Utility is a function of (among other things) the focus of the evaluation, and particularly of whether it responds to UNICEF’s needs by filling gaps in evidence on important operational priorities. Delivering a technically proficient evaluation that reproduces existing knowledge delivers less utility than delivering a technically proficient evaluation on a subject where important uncertainties lie and where answering the evaluation questions, will unlock value for the organization.

The construction of the portfolio of evaluative activities involves a complex composite of centralized and decentralized mechanisms. The 2018 *UNICEF Procedure on the Implementation of the 2018 UNICEF Evaluation Policy* establishes benchmarks for evaluation coverage at the country, regional and corporate levels, along with responsibilities for evaluation management and quality assurance. Heads of Offices are responsible for meeting the assigned benchmarks on coverage and on:

- a. Conducting at least one country thematic evaluation per year by each country office, except for small offices, which may conduct three such evaluations per programme cycle

- b. Country programme evaluations managed by regional offices at least once every two country programme cycles or once per programme cycle if monitoring and audit information points to a significant shift in the programming context or a significant increase in the level of risk
- c. Evaluation of humanitarian response by various levels of the organization as specified in the *Evaluation Policy*
- d. Multi-country regional and corporate evaluations by regional offices and the Evaluation Office, respectively

In itself, this list, while clear and easy to understand, does not explicitly draw attention to evaluating gaps in the existing evidence base and relevance to the operational success (though this is not to say that these issues will not be touched on through application of the benchmarks or shape decision-making less formally). It also leaves open the question of the representativeness of the overall balance of work and proportionality. At the HQ level, there was a view, with which we agree, that there is a need for a better system for identifying what should be evaluated. Overall, a strengthened and more systematic focus on plugging evidence gaps relevant to the achievement of operational objectives is needed and could be achieved via a specific project to map evidence gaps and how the UNICEF evaluation function can be marshalled to addressing them. It is very welcome that such a project is already underway, due to be completed by March 2023. A similar mapping of gaps at the UNICEF Strategic Plan level has been undertaken for research.²⁴

Pursuing incremental improvements in shaping the direction of evaluative efforts to strengthen utility can, though, go hand in hand with maintaining sensitivity to existing processes at country and regional levels. A more utilization-focused approach to portfolio planning could support a portfolio of evaluations with greater relevance to operational and strategic priorities. Overall, UNICEF has taken important steps to improve its selection of evaluation topics.

Improving the planning and management of evaluations, including through strengthening stakeholder engagement

While evaluation topic selection is vital to achieving utility and should also be a function of independent direction, how evaluations are conceptualized and then planned and managed can be equally important. A document providing guidance for UNICEF's evaluation practice from start to finish does not yet exist, though there is guidance covering the steps along the way. This includes guidance on, for example, developing the Terms of Reference for an evaluation and guidance on reporting. The 2017 *UNICEF-Adapted UNEG Quality Checklist for Evaluation Terms of Reference* provides detailed guidance for this critical step, outlining the purpose, scope, process and products for an evaluation and clarifying expectations (in this respect, complementing the GEROS quality criteria). The document establishes the key dimensions to be included in the Terms of Reference and encourages opportunities for stakeholder engagement.

There is, however, little formal guidance on the 'upstream' process of first conceptualizing the need for an evaluation, including how to explore the ways different stakeholders articulate their needs (for example, through the systematic use of sandpits or other multi-stakeholder forums). These kinds of structured consultations, to enable stakeholders to inform the commissioning process, would not require a whole new system but could be better incorporated into existing practices. Early and structured engagement with stakeholders could help UNICEF evaluations by: defining and

²⁴ [MegaMap on Child Well-being Interventions in LMIC's and associated products \(unicef-irc.org\)](https://www.unicef-irc.org/publications/megamap-on-child-well-being-interventions-in-lmic-s-and-associated-products)

scoping the issue; agreeing on a common language and terminology; sharing participants' expertise and understanding of the problem; using creative and innovative thinking techniques to focus on a problem; and forging evaluation questions, evaluation frameworks and draft Terms of Reference.²⁵

Similar processes used towards the end of the evaluation process can help to engage stakeholders in translating findings into relevant key messages and actionable recommendations for evaluation users. Several UNICEF and external stakeholders consulted for this peer review noted the benefits of holding workshops on evaluation recommendations, which can help evaluators arrive at more feasible, acceptable and suitable recommendations. Engaging in such 'recommendation co-development' processes more frequently, and framing them as a standard, expected step within evaluation terms of reference, could help to strengthen UNICEF's decentralized structures.

Ensuring the timeliness of evaluations

A key issue raised by several internal stakeholders was that evaluations, particularly decentralized evaluations of regional and country programming, are not always carried out in a timely manner. One of the interconnected issues here is also linked to timely and predictable funding for evaluations and the ability of the decentralized evaluation function to initiate them on time. Evaluation design should go hand in hand with programme design, not least so that programmes can collect the necessary monitoring data to inform assessment of the contribution of programmes to outcomes. Currently, this does not happen consistently and systematically in all cases, but will only become more important as UNICEF moves towards greater use of impact evaluation, which requires sufficient pre-intervention data as well as data collected across time.

Utility is also manifested by the enabling environment for evaluation, through timely use of learning from evaluation within decision-making processes, but the timing of evaluations does not always cohere with the cyclical schedule of regional- and country-level planning and decision-making. To increase the utility of evaluation, there is a need to ensure that evaluations are planned and happen at the right time – and in time for the lessons and recommendations they produce to be applied by stakeholders in planning decisions for the next programme cycle. As suggested in the 2018 *Evaluation Policy*, the uses to which evaluations will be put should already be at the forefront of evaluation managers' minds at the evaluation planning and design stage; our consultations with internal stakeholders suggest that this leaves room for improvement.

Improving the format and accessibility of evaluation products

The format of evaluation products was raised by several stakeholders as limiting their optimal utility. While, as noted above, reports rated by Geros as 'satisfactory' can suffer from inappropriate formatting and length, it was also clear from discussions with UNICEF staff and external stakeholders (particularly from programme and donor governments that we engaged with) that there is demand for shorter, more digestible and accessible evaluation products, with clear, relevant, actionable messages. There is a place for comprehensive evaluation reports, but there is also a clear need for a more varied set of communication tools. Short reports and summaries, presentations and infographics could be used to better effect to disseminate the top 'take-aways' for intended audiences. It would be beneficial to use Terms of Reference to set requirements for a range of deliverables, including standard reports but also more easily useable products designed to maximize accessibility to intended evaluation users.

²⁵ See, for example, <https://petranetwork.org/wp-content/uploads/2020/05/Sandpit-methodology-overview-March-2020.pdf>

Another useful area to consider would be how evaluation products are presented in the public arena, particularly through the UNICEF and other UN websites. A more user-friendly online library of evidence would also help to improve the visibility and accessibility – and therefore the utility to key stakeholders ranging from UNICEF managers and staff to the wider community of interested policymakers and researchers – of the very considerable range of evaluation products currently produced but often under-utilized.

Formal compliance with management response process

Part of UNICEF’s approach to supporting utility that has come into much greater focus in recent years is the management response to an evaluation. We discuss the relationship between management responses and utilization of evaluations further in the section below on ‘Follow-up and use’ and so discuss it only briefly here. The key point in relation to the utility of evaluations (as opposed to their actual utilization) is that, while there is a very high level of formal compliance with management responses (as many as 99 per cent of evaluations conducted between 2018 and 2020 have a management response), the level of compliance is not a reliable indicator of the utility of evaluations. Ascertaining whether evaluations are genuinely useful and meaningful can be in part understood through discussions with stakeholders, for example, on the relevance of the topics selected, and understanding of the decisions that evaluations may have influenced.

Roles and responsibilities

Key findings:

While most roles, responsibilities and accountabilities for evaluation are clearly set out in the *2018 Evaluation Policy*, there are some areas in need of clarification in the coming policy.

Current roles and responsibilities for evaluation, particularly at the decentralized levels, do not always guarantee the independence of the evaluation function or consistently support the wider aims of the *2018 Evaluation Policy* at decentralized levels.

Mechanisms for enforcing accountability for evaluation, including reporting and monitoring, currently result in variation in how roles are fulfilled throughout the function, particularly at decentralized levels, and should be strengthened.

Alternative potential options each carry their own potential advantages and challenges.

Formal roles and responsibilities

Roles and responsibilities for the evaluation function are articulated in the *2018 UNICEF Evaluation Policy* and summarized in Annex 6 of this report. These include the respective roles and responsibilities within the function itself, including the roles of the Evaluation Office at global level and the REAs and (M)CES at decentralized level. They also include the roles and responsibilities of other actors outside the evaluation function.

Accountability arrangements at HQ and decentralized levels

The practitioner’s views, as revealed in the survey data, concerning accountability arrangements at HQ are relatively clear: most think they are either somewhat or fully adequate at all three levels of the organization. However, the same data suggests that some 10 to 15 per cent of respondents view accountability arrangements at the regional and country levels as ‘not at all adequate’ or ‘not very adequate’. Notably, only just under half of practitioners reported that roles and responsibilities for data, research and knowledge management regarding evaluation are defined clearly in their

office/division. Practitioners also indicate that the most helpful things in delivering the *Evaluation Policy* would be clearly defined responsibilities and evaluation norms and targets.

Senior management reported perceptions of much lower levels of adequacy in how the *Evaluation Policy* assigns accountabilities, with only 56 per cent reporting that accountabilities at HQ level are fully or somewhat adequate, 60 per cent reporting adequacy of accountabilities at the regional level and 62 per cent reporting adequacy of accountabilities at the country level.

Figure 9: Practitioners’ survey: To what extent is the 2018 Evaluation Policy adequate in assigning evaluation accountabilities at different levels?

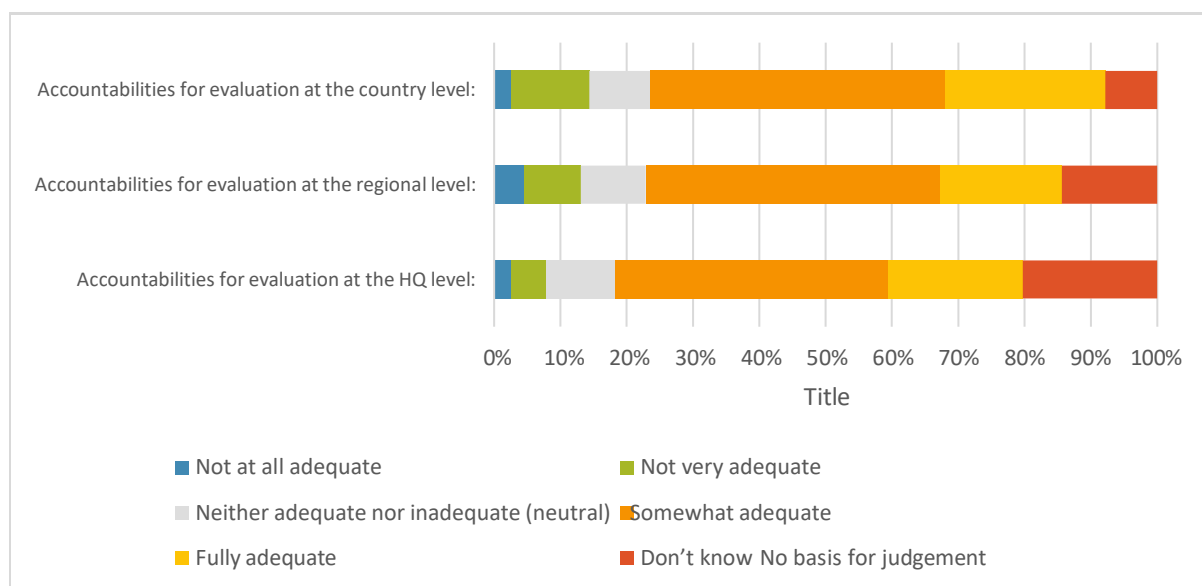
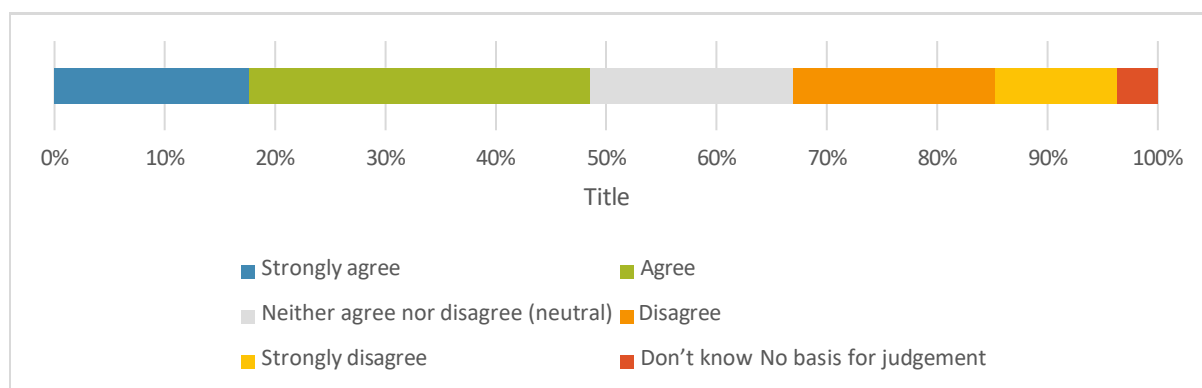


Figure 10: Practitioners’ survey: To what extent are the roles and responsibilities for data, research and knowledge management vis-à-vis evaluation defined clearly in your office/division?



Roles, responsibilities and accountabilities carry risks to independence

The formal roles and responsibilities for evaluation in UNICEF may not adequately guarantee the independence of the evaluation function particularly at decentralized levels, where regional- and country-level evaluation staff report primarily within the offices at which they are based, rather than directly reporting up through the evaluation function. Within the UNICEF decentralized evaluation structure, independence may be affected by priorities of staff who sit outside the evaluation function. Consequently, the independence and quality of evaluation may be exposed to the preferences of proximate (regional- or country-level) management.

The 2018 *Evaluation Policy* notes that “for the regional evaluation adviser, the regional director provides management oversight and guidance, and the Director of Evaluation provides technical guidance”, but the mechanisms listed to give the Director of Evaluation effective organizational traction are weak (especially when compared with the more substantive reporting lines to Regional Directors). This creates a structure wherein REAs are subordinate to, and therefore not independent of, the RDs, which can play out differently in different regions. The 2023 *Revised Policy* should address this, considering the merit of potential alternative options set out below. There should at minimum be a reconsideration of the tasks carried out by REAs that are reported to the Director of Evaluation, with serious thought given to whether the substantive tasks, including monitoring of Key Performance Indicators (KPIs) and Costed Evaluation Plan (CEP) implementation, should be conducted under the direct supervision of the Director of Evaluation. Similarly, there were concerns about country evaluation staff reporting to country representatives, which may reduce their ability to retain their independence, suggesting it may also be worth considering how to strengthen reporting lines of these staff to achieve balance in meeting their independent evaluation role in ways that supporting accountability and learning for country offices.

Mechanisms for ensuring accountability in relation to funding, establishing adequate capacity and applying norms and standards could be strengthened

While most roles, responsibilities and accountabilities for evaluation are clearly set out in the 2018 *Evaluation Policy*, there are some areas in need of clarification. Importantly, mechanisms for enforcing accountability, including reporting and monitoring, are weak, which results in variation in how roles are fulfilled throughout the function, particularly at decentralized levels.

In relation to allocating and spending sufficient and appropriate funds for evaluation, particularly at the decentralized level, the peer review panel noted a risk in accountabilities. According to the 2018 *Evaluation Policy*, Regional Directors should ensure adequate capacity development and the allocation of adequate funds by representatives to the evaluation function. The Evaluation Office appears to be accountable for the oversight of this but has to date apparently not received monitoring reports on the allocation of resources to evaluation from the regional level. Consequently, monitoring and accountability for the allocation and spending of sufficient and appropriate funds for evaluation at regional levels appear weak or lack visibility.

The 2018 *Evaluation Policy* asserts that country representatives should also provide adequate financial and human resources for the implementation of the CEPs. Our understanding is that this is not always followed through and enforced, in the sense that there is no ‘ringfencing’ of resources for CEP work. If REAs monitor this, we are told that this is not consistently reported up to Evaluation Office. This appears to have different implications for evaluation in different regions. At CO level, resource allocation for evaluation should reflect the aims set out in the CEP. Furthermore (as was raised by some stakeholders), because COs usually prepare CEPs with resources that are not ‘ringfenced’, CEPs are not monitored in terms of budget, and budgets are continuously adjusted, thus making it harder to enforce the CEP expenditure and making the evaluation function at CO level harder to deliver efficiently.

This is an area that should require particular attention in the *Revised Evaluation Policy* for 2023. There appears to be an inherent weakness in evaluation monitoring and reporting. For example, under the revised policy, Regional Directors were provided with additional human resources (Regional Evaluation Advisors and monitoring specialists) whom they directly supervise and were granted US\$2 million each from the Evaluation Pool Fund in 2019 and more funds in subsequent years. The impact of these new capacities at regional level could be traced to the increased number

of evaluations. However, it was difficult to understand whether the financial resources allocated to regions have been used strategically. The total additional investment in evaluation in the subsequent year could be traced to the Evaluation Pooled Fund allocation, meaning that there was little or no multiplier effect due to the Fund. The issue here is a lack of any obligation to monitor and report on the use of additional resources by the recipients.

The uneven consequences of the *2018 Evaluation Policy* can also be seen in relation to establishing adequate capacity for evaluation management in country offices. The language used in the policy is more of an encouragement than a hard requirement: “In order to ensure dedicated evaluation positions, country offices may decide (in large programmes) to establish an evaluation specialist post, while smaller offices may pool resources with other UNICEF offices to fund a multi-country post dedicated to evaluation.” An incremental improvement would be to require that all COs over a certain size (say, US\$100 million) recruit a person of a certain seniority (say, P4). This would bring the seniority and weight of the evaluation function more into line with other UNICEF functions.

Furthermore, the policy allocates responsibility to country office evaluation specialists for ensuring the application of norms and standards for country-level evaluations. It is not clear from the policy how these responsibilities are to be reported and monitored, nor is it clear what the role of specialists should be in such monitoring. Experiences may vary, but it was suggested to the peer review team that, at least in some countries, this resulted in specialists having responsibility without the resources needed to exercise this responsibility. The lack of control or guaranteed influence over the country evaluation spend may also limit the independence of country evaluation staff.

Potential options for formal roles and responsibilities

It would be beyond the scope of this peer review to make prescriptive recommendations to change reporting lines, which are rather a decision for UNICEF to take based on its organizational priorities. However, we can point to a range of potential options for formal roles and responsibilities to ensure and protect independence, along with their key potential advantages and challenges. These are set out in brief below.

Status Quo: The Evaluation Director reports functionally and administratively to the Executive Director. The Evaluation Director reports at least annually to the Executive Board. Protections of independence are given through provisions in the *Evaluation Policy*. Reporting to the Oversight Committee provides mechanism to raise issues of concern.

Potential Advantages: The Evaluation Director can raise issues on independence to the attention of the Executive Board in the annual session, while maintaining connection to the organization as a member of the senior management team. Connection to the organization may help to be better attuned to evidence needs, gaps and challenges, which may contribute to better use and relevance of evaluations.

Potential Challenges: This may not achieve the highest level of real or perceived independence. There may be greater potential risk for real or perceived management influence.

Alternative option 1: A fully independent Evaluation Office. The functional and administrative reporting line of the Director Evaluation Office is changed from the Executive Director to the Executive Board.

Potential Advantages: This reduces the potential risk of real or perceived management influence.

Potential Challenges: The Executive Board is not currently set up to provide effective day-to-day oversight of the UNICEF evaluation function.). Notably, there is no Board sub-unit specifically set up

to focus on oversight issues. A potential opportunity is that the UNICEF Executive Board, together with the boards of UNDP, UNFPA, UNOPS and UN-Women, is commissioning a review of their effective discharge of the Executive Board's oversight role, so there may be opportunity to mitigate some of the potential management challenges.

Alternative option 2: Align Regional Evaluation Specialists to report directly to the Evaluation Office (for example, the Evaluation Director or Deputy) while maintaining a dotted line to the Regional Directors.

Potential Advantages: This may reduce potential risk of real or perceived management influence. There may be potentially stronger control over execution of decentralized evaluation workplan.

Potential Challenges: It may be a challenging role for regional evaluation specialists to play, being situated in a regional office and seeking to be attuned to the issues and needs of the regional office, but taking instruction from Evaluation Office.

Follow-up on evaluations and use

Key findings:

- There is a commendably high level of formal compliance with management response processes, but this does not necessarily indicate that evaluations are well used by the organization as a whole or by individual programmes.
- The Evaluation Management Response process is helpful for monitoring purposes but could be routinely supplemented by engagement with evaluation users to achieve a deeper understanding of what they value most highly and what could be improved.
- There is potential to create a more user-friendly digital repository of evidence and to promote this among UNICEF and other UN staff, as well as external policymakers and researchers, which would in turn help to secure UNICEF's position as leaders in generating and organizing evidence to improve the situation of children.
- There is potential for improvement in the evaluation communications function to craft effective messages communicating the evaluation findings and implications for target audiences, to increase the visibility and enable broader use of findings.

Very high formal compliance with management response processes does not necessarily indicate that evaluations are well used within decision-making

Management responses are a central part of delivering an evaluation function that has utility. Supporting the management response is an Evaluation Management Response (EMR) system.²⁶ EMR data track responses and are reported to the Board, the Global Evaluation Committee and the Regional Management Teams. According to the 2021 UNICEF evaluation function annual report, "Management responses are an essential platform for ensuring the implementation of evaluation recommendations. The UNICEF *Evaluation Policy* requires that management responses be submitted within 60 days following the posting of an evaluation in the Evidence Information Systems Integration (EISI) database. While the time frame was extended in 2020 and 2021 to 90 days due to the COVID-19 pandemic, it was reinstated to 60 days in 2022. Overall, UNICEF offices are submitting

²⁶ UNICEF Evaluation Office. 2018. Evaluation Management Response, Guidance for UNICEF Staff. UNICEF. As of 8 December 2022: <https://www.unicef.org/media/54801/file>

management responses on time. As many as 99 per cent of evaluations conducted between 2018 and 2020 have a management response.”

While there is an impressively high level of formal compliance with management response processes, this does not necessarily mean evaluation evidence is driving decision-making among programme teams. As reported above in the section on utility, around three quarters of senior management reported that they found evaluation useful or very useful, while less than two thirds of practitioners reported that evaluation products had been effective in meeting operational and strategic decision-making purposes in their office or division. Some stakeholders, including from among the UNICEF Executive Board, have expressed the concern that management responses do not always commit to specific actions that can be followed up and monitored, which can undermine confidence that programme teams are responding sufficiently to evidence. To build a more systematic understanding of the extent of the use of evidence in programming, it may be useful to conduct a corporate evaluation similar to that completed recently by the German Development Agency.²⁷

Going back to 2016, UNEG’s research *Evaluation Use in the UN System: Conclusions from the Data* noted, “ eported levels of evaluation use vary within and across agencies; evaluation use does not appear to be systematic across the board. Many respondents shared anecdotal evidence of use, but did not consider that use in their agencies was high overall.” Improving use of evaluation findings has therefore proved to be a long-lasting challenge (and not only among UNICEF, as is apparent from our light-touch review of other agencies). As identified by the UNICEF Evaluation Office’s 2018 report *Influential Evaluations: A Selection of UNICEF Evaluations that Led to Learning and Change*, delivering more useful evaluations “requires a shift away from viewing evaluation only as an accountability exercise, but also as an opportunity to leverage credible evidence for children.” As noted earlier on in this report, we very much welcome the study currently underway to assess operationally important evidence gaps and encourage this to be used by the Evaluation Office to steer the evaluation portfolio towards those areas that will be most useful in supporting organizational learning and programme improvements.

Further, while the Evaluation Management Response process is helpful for monitoring purposes, this could be routinely supplemented by engagement with evaluation users to achieve a deeper understanding of what they value most highly and, conversely, could be improved. This could be done through follow-up meetings or specific sessions dedicated to the topic.

There is potential for greater collaboration across UNICEF knowledge functions to improve the wider visibility, and so support the use, of evaluations

Strengthening collaboration with other UNICEF knowledge functions would help to bring some further benefits for the utilization of evaluation evidence. To take two examples, as discussed above in the section on utility, the development of a more user-friendly online library of evidence could help to improve the visibility, accessibility and therefore use of evaluations among not only UNICEF staff but the whole community of interest engaged in the work of improving the lives of children. This would in turn help to further secure UNICEF’s position as leaders in generating and organizing evidence to improve children’s lives.

Second, there is significant potential for more collaborative work with, and leverage of the expertise of, the communications function, to craft effective messages communicating evaluation evidence

²⁷ [How evidence-informed is German Development Cooperation? | DEval - Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit GmbH](#)

and implications for target audiences. This, again, could help to increase visibility and enable broader use of findings. Our consultations with internal stakeholders suggested there is limited and inconsistent liaison between the evaluation and communication functions, which prompts us to emphasize the potential to nurture a much more mutually beneficial relationship: while the communication function would benefit from ‘mining’ evaluations for powerful content, the evaluation function would benefit from more effective publicity to promote learning and engagement with the lessons it produces.

These steps cannot, of course, be achieved by the evaluation function alone but depend on the interest and willingness of other functions to support efforts to strengthen the utilization of evaluation evidence.

The enabling environment

Key findings:

- Modalities of funding, including the irregularity of when and how much funding is allocated to evaluation at different levels, do not currently consistently support effective planning and implementation of evaluations. It is management’s responsibility to ensure sufficient and predictable funds are provided for evaluation.
- The question of the sufficiency of the 1 per cent allocation is challenging to answer with the available evidence. The peer review learned of concerns regarding its calculation.
- The evaluation theory of change is welcome but could be better used as a living document and be further developed to explain causal pathways, with greater attention to what is within the evaluation function’s sphere of control and what is within its sphere of influence.
- UNICEF is driven by laudable goals and the evaluation function struggles at times for organizational attention in this environment. Strengthening UNICEF’s evaluation culture outside the evaluation function would help to ensure lessons on what works are valued as fundamental to the mission of improving the situation of children and applied within decision-making. This requires action not only from the evaluation function but also from UNICEF leadership.
- More could be done to maximize the benefits of evaluation through collaboration with other functions (including ethics, data, innovation and communication functions), which is not currently routine. We support the Evaluation Director’s vision in this regard.
- In particular, linking evaluation with digital transformation is a substantial challenge for the future but is essential to ensuring that UNICEF can leverage the advantages of digitally enabled evaluation methods.

Ensuring funding and funding modalities support effective evaluation

The 2018 *Evaluation Policy* establishes a target of 1 per cent of expenditure to be allocated to the evaluation function. The *Evaluation Policy* states, “Globally, to meet minimum evaluation coverage in accordance with the *Evaluation Policy*, UNICEF will allocate at least 1 per cent of its overall programme expenditure to evaluation. Planning studies, monitoring, surveys and research must be financed separately from evaluation.” Board approval for the Evaluation Office budget is also required under current policy, which states, “As part of the approval of the integrated budget, the Board approves the budget of the Evaluation Office.” This apparently straightforward position is complicated by at least three questions: first, the definition of the 1 per cent; second, the ways in

which funding is made available ('modalities of funding'); and, third, whether the 1 per cent is sufficient and sufficiently predictable. We consider each of these issues in turn.

Defining 1 per cent

While there has been a formula used for several years to define the 1 per cent, it is unclear to those responsible for overseeing the evaluation function how it was established. We were told that, in its present form, the numerator and denominator use different metrics. The numerator expresses evaluation expenditure as consisting of both actual expenditure and spending commitments, while the denominator is expressed only as the actual programme expenditure. Using this definition, the 2021 figure amounted to 0.91 per cent. The Evaluation Office estimates, however, that, if both the numerator and denominator includes actuals and commitments, the percentage for 2021 would be 0.71 per cent. Alternatively, using evaluation actuals as a percentage of programme actuals, the Evaluation Office estimates that the percentage in 2021 is 0.53 per cent.²⁸ There is a concern, then, that using the current different metrics for the numerator and denominator may skew the picture of how much is being spent on evaluation.

Further, we were told that there is an ongoing and complicating question about what precisely an 'evidence function' is and how this relates to the 'evaluation function'.

Overall, our consultations with UNICEF stakeholders found confusion about what the 1 per cent figure for expenditure means in practice (with some indicating that it was a pot created for all evaluations while others believing it was for each programme to spend 1 per cent, and still others uncertain what it was 1 per cent of).

Modalities of funding

In the section on 'Roles and responsibilities' above, we describe the challenges of monitoring, accountabilities and lines of reporting at decentralized levels. Within this system, there is a lack of accountability for implementation of the 1 per cent policy, and so, while evaluation spending at regional and country levels could theoretically be sufficiently well targeted and predictable to ensure the efficient and effective use of funds, there are weak mechanisms for ensuring that this is always the case. Accountability and reporting arrangements would need to be strengthened for there to be confidence that the 1 per cent is systematically achieved.

Relatedly, currently at the HQ evaluation level, some funding is delivered quarterly in tranches and with uncertain amounts. At the decentralized level, with the Evaluation Pooled Fund, the Evaluation Office is responsible for reporting its use and effectiveness, but the regional office has authority over how the funding is spent. The expectation is that HQ and regional offices will maintain a dialogue on allocation criteria and on specific country office allocations and rationale. Country offices are expected to be transparent in their use of funds and the justification for this along with reporting on the country thematic funding. Country representatives are encouraged to spend the allocated 1 per cent, but this can be challenging for them to achieve when longer-term planning is undermined by unpredictability of evaluation funding. Equally uncertain is the extent to which evaluation topics are selected on the basis of what would contribute most to delivering the *Evaluation Policy*.

The Evaluation Office is consequently not confident that it is able to ensure that the significant potential benefits of thematic pooled funds are achieved in practice. Key potential benefits have been identified as: supporting strategic funding of programmes by increasing flexibility in the allocation of resources to areas of highest programme needs; providing a more flexible, longer and

²⁸ PowerPoint slide set entitled 'Resourcing of the UNICEF Evaluation Function: Trends, Projections, and Current Gaps' dated August 2022 and made available to the review team by the Evaluation Office

harmonized timespan for using contributions, which also helps to reduce transaction costs; allocating sufficient funds to strengthen equity-focused programming and effective implementation and monitoring; channelling humanitarian funds in a more flexible manner to respond to Core Commitments for Children in Humanitarian Action and promoting initiatives that prioritize strengthening the ‘nexus approach’.²⁹ On its own, the Evaluation Office is unable to resolve this funding modality issue, which would require action by UNICEF management or at Executive Board level.

Is 1 per cent sufficient?

Definitional disputes aside, and based on the existing formula, we can address the sufficiency of the budget in one of two ways. The first is to compare the percentage set aside with other UN agencies. The second is to understand what is not being delivered in terms of evaluation, but that should be, as a result of insufficient funding and what impact and value this absence has.

On the first question, a brief review was made available by the UNICEF EO suggesting a range of percentages set aside for the evaluation function of other agencies, as described in the table below

Table 1: Evaluation target funding in a sample of other UN agencies

Agency	Target percentage for evaluation
UN Women	2-3% of total programme expenditure
FAO	0.8% of the regular programme budget
WFP	0.4% total contribution income
UNESCO	3% of programme expenditure
IFAD	0.90% of the programme of loans and grants
OIOS	0.5-3% of entity expenditure
UNHCR	No percent
ILO	2% of project expenditure
UNFPA	1.4-3% total programme expenditure

This provides a broad sense of the range of set-aside funds. It should be noted that different operational characteristics may require different levels of evaluation and, in general, these numbers should be treated as only broadly indicative. Nonetheless, within this range, different formulae are used, but the numbers suggest that UNICEF is towards the median to bottom end of the range of target funding for evaluation functions.

Estimating sufficiency based on the second criterion is problematic. There is no metric defining what is actually delivered through the evaluation function, as opposed to what would be functionally optimal to support operations, or what is wanted by donors, management and other stakeholders to ensure accountability and learning. Without this, or at least some estimation of the scale of the issue, it is not possible to be confident about what would happen if the current budget were fully spent or were increased. However, the Evaluation Office has suggested³⁰ that additional funding would facilitate more rigorous and expensive evaluation work in relation to: National Evaluation Capacity Development; evaluations at the outcome and impact levels; innovations in evaluation; joint evaluations with other UN agencies; institutional effectiveness evaluations; and humanitarian action evaluations. It would also facilitate more impact evaluations (which, as noted above, can be

²⁹ See the draft *2022-2025 Guidelines for Allocation of Thematic Funds*.

³⁰ PowerPoint slide set entitled ‘Resourcing of the UNICEF Evaluation Function: Trends, Projections, and Current Gaps’ dated August 2022 and made available to the review team by the Evaluation Office

expensive and resource-intensive). Currently, the EO estimates that UNICEF conducts one impact evaluation for every US\$900 million spent on programmes.

The evaluation theory of change is welcome, but some elements could be strengthened to better support developments in policy and practice

As envisaged in the 2018 *Evaluation Policy*, a theory of change for the UNICEF evaluation functions was developed that sets out its intended approaches, drivers and inputs, outputs, outcomes, impacts, and assumptions and risks (see Annex 9). A theory of change is not intended to be a detailed management plan illustrating every causal pathway, but it is noticeable from the theory of change; if the policy is to deliver the intended impacts, much will depend upon mechanisms of change that are currently unstated. There are (implicit) logical and credible links between the stated outputs, outcomes and impacts, but our consultations with UNICEF stakeholders indicated that they are neither fully agreed upon nor understood throughout the organization. To be maximally useful, a theory of change in a rapidly changing environment should be revisited, adapted and used as a 'live' tool to inform discussion of changes in policy and practice. This does not appear to have happened very frequently or consistently across the various levels of the evaluation function, nor in other functions.

Relatedly, it may be helpful for the theory of change to distinguish between those aspects of evaluation policy, processes and practices, over which the UNICEF evaluation function has direct control, and those that require the engagement and collaboration of other UNICEF functions and indeed external organizations. This may help emphasize how important it is that the evaluation function needs management collaboration and partnership to influence the wider enabling environment for evaluation. The next evaluation policy (and future iterations of the theory of change) might focus more explicitly on building the capacity for partnership working within UNICEF, across UN agencies, and with the wider external enabling environment.

Strengthening UNICEF's evaluation culture outside evaluation practitioners

An organization with a culture of evaluation is an organization that actively seeks out and applies evaluation evidence, regards the evaluation function as 'everyone's business' and builds links between evaluation and other knowledge-generating functions. An organization intent on strengthening the culture of evaluation would have a clear strategy for achieving this.

The evidence gathered for this peer review suggests that UNICEF has an uneven evaluation culture, with inconsistent use of evaluation evidence and insight. When it is absent, little is done to challenge this. Sometimes proudly, and sometimes sadly, we were told often that UNICEF is a doing, not a thinking organization. We identified across various different functions, levels and geographic regions, a less-than-enthusiastic view of evaluation among many non-evaluation staff (who may view it as a sanction or means to expose poor performance rather than as an opportunity to learn and improve), while evaluation practitioners often sense a lack of demand for their work. The weakness of demand for evaluation was said by some to undermine the function and UNICEF's overall ability to meet its potential to become a leading knowledge broker for children.

However, there may be a latent (or 'pent-up') demand for evaluation that can be stimulated through different approaches to engagement and dissemination of learning, and more diversified evaluation products. Building demand would involve showcasing work that is timely and makes a difference. As suggested in the section on utility, shorter and more accessible evaluation products may support better engagement with key lessons from the evidence. Nonetheless, it is clear that, while the evaluation function itself has a vital role to play in further promoting and embedding a culture of

evaluation and learning, UNICEF leadership and management in other functions have an equally important role in setting the tone and vision of the organization as a whole. Evaluation culture cannot be confined to the evaluation function and, while the function may – and should – seek to promote the value of evidence and learning, some of the factors affecting this (including the organization’s leadership and other functions) are outside its sphere of control.

Maximizing the benefits of evaluation through collaboration with other functions

UNICEF has enormous potential to leverage its very significant resources, expertise and global reach to maximize the benefits of evaluation. While collaborative working across functions is key to unlocking these benefits, the findings of this peer review suggest that there are a number of gaps in whole system working at UNICEF, which represent missed opportunities to achieve its mandate for children.

Evaluation Office stakeholders noted that relationships between the Office of Research and the Evaluation Office are strong, with several current EO staff having previously worked at the Office of Research or having research backgrounds. These kinds of strong cross-function relationships are an essential part of ensuring the relevance and value of evaluation to large, multilateral organizations.

As suggested earlier in this report, a further area for fruitful and mutually beneficial collaboration would be with the communication function (while evaluation can provide powerful content for communication, effective communication is crucial to get key messages from evaluation out to target audiences). Similarly, the relationship between evaluation function and data and monitoring should support more efficiency at country level with more streamlined and prioritized demands. The Organizational Design and Improvement Lab (ODIL), for example, aims to support agility and responsiveness through programme acceleration, business innovation, and better oversight, coordination and monitoring. There thus appears to be clear potential for greater synergy with the evaluation function. We do not suggest that the aim should be a homogenized, fully integrated knowledge function, but rather that greater connection across knowledge (and other relevant) functions, with clarity of purpose, would bring benefits to each. These issues might usefully be addressed within the 2023 policy review. By grasping the opportunity to link its knowledge and learning functions more coherently, UNICEF could transform its ability to learn, adapt and contribute to global understanding about how the rights of children might best be met.

Linking evaluation with digital transformation

UNICEF is moving at speed to support the digitalization of its functions. There is an expressed interest to engage in UNICEF’s digital transformation from both the Director of Evaluation and the Chief Information Officer. The 2023 *Evaluation Policy* is a timely opportunity to strengthen the value of the evaluation function in shaping the opportunities and meeting the challenges of digital transformation.

UNICEF’s digital transformation is relevant for evaluation in several ways. First, the digitalization of UNICEF’s services (from e-health to e-learning and beyond) needs to be fully evaluated. Second, digitalization introduces new opportunities and risks for data collection and analysis, which need to be better understood. Third, there are substantial opportunities for transforming knowledge management (particularly with regard to the role of AI in creating effective search functions). This is already envisaged in the evaluation plan 2022-2025, which promotes the idea of UNICEF innovation labs using new sources of data and conducting impact evaluations that use data in novel ways.

Further, digitalization raises many important evaluative questions regarding the digital divide, child protection and data privacy, and the ethical and impact dimensions of online learning and digital

health systems. Wider evaluative questions concern the effectiveness, efficiency, impact and implications of ‘frontier technologies’, such as block chain and crypto currencies, which are already impacting on children and on the services UNICEF provides. Real-time data and combining data from multiple sources, while protecting rights to privacy, open up radical new opportunities and challenges for the evaluation function.

The 2021 annual report on evaluation states, in relation to enhancing the use of evaluation in innovation, that Heads of Offices should plan for the use of evaluations even as they are being designed. It further states that the Director of Evaluation, working with Heads of Offices, should invest in innovative technologies to advance the use of evaluations, including platforms for connecting evaluators with users of evaluation evidence. This could work comfortably with the wider use of platforms for structured stakeholder engagement discussed above in the section on utility.

The external enabling environment for evaluation

According to its website, “UNICEF is the UN organization mandated to protect the rights of every child, everywhere, especially the most disadvantaged, and is the only organization specifically named in the Convention on the Rights of the Child as a source of expert assistance and advice.”³¹ This mandate to provide expertise and advice highlights how crucial the evaluation function is to global efforts to improve outcomes for children through evidence-led policy and practice.

UNICEF already contributes to this global environment. For example, various methods and approaches are communicated in a series of UNICEF webinars and e-learning opportunities. Other important examples of valuable contributions to global evidence include, to take just a few examples: its global evaluations on education alongside the special issue of the International Journal of Educational Development;³² the global review of influenza vaccine introduction conducted with WHO, which focused on the education response to COVID and leveraged UNICEF’s access to data in 153 countries;³³ and the synthesis of Rohingya response evaluations (commissioned with the International Organization for Migration and UNHCR).³⁴ These (and other) examples demonstrate UNICEF’s capacity for global reach in the evaluation ecosystem.

Looking forward, the *Plan for Global Evaluations 2022-2025* includes a commitment to give greater emphasis to supporting national evaluation capacity and states that, by strengthening its own ‘whole-of-evaluation-function’, it will contribute to global evaluation capacity. The plan also commits to increased joint and inter-agency evaluation, where appropriate, and strengthening system-wide evaluations. It commits to forming a methods innovation lab that would not only service the immediate needs of the evaluation function in UNICEF but also contribute global evaluation capacity. The plan also states, “The Evaluation Office will continue to co-lead EVALSDGs and serve as an active member of the EvalPartners’ Management Group. Through these partnerships and others, UNICEF will continue to support the preparation of voluntary national reviews.” The plan aims to support in-country capacity and joint funding of evaluations and to strengthen the design of experimental, quasi-experimental and non-experimental impact evaluations. Further, it is welcome that HQ evaluation staff signalled interest in supporting the Global Evaluation Initiative to promote national evaluation capacities.

³¹ <https://www.unicef.org/child-rights-convention/unicef-role>

³² <https://doi.org/10.1016/j.ijedudev.2021.102485>

³³ <https://www.tandfonline.com/doi/full/10.1080/14760584.2019.1640119>

³⁴ <https://www.unhcr.org/5e453ea64.pdf>

These are appropriate aspirations for UNICEF's evaluation function, which has a uniquely important contribution to make to the understanding of what works in protecting children's rights. While recognizing the institutional arrangements that risk achievement of this goal – particularly competition for management time and resources – we urge that this work contributing to the external enabling environment receive the priority it deserves.

Conclusions

The strategic question of the peer review panel was:

Based on the experience implementing the 2018 Evaluation Policy for UNICEF over the past five years, what aspects of the policy have worked well in practice in optimizing the evaluation function in the organization in accordance with the UNEG norms and standards and should therefore be retained and potentially built on, which aspects have not worked well in and should be changed or abandoned, and, in retrospect, what gaps evident in the 2018 policy need to be filled in the 2023 Evaluation Policy?

There is no simple answer to this question, but there is a relatively clear picture of overall progress.

Our key conclusions are:

- Progress towards delivering evaluation policy aims is broadly good with some areas for enhancement. Policies, guidelines, data on evaluation quality, management responses and the existence of a specific funding target for evaluations, are areas of strength in UNICEF's evaluation function. This is evidenced in the survey responses from evaluation practitioners and management, from GEROS data, from the review of evaluations rated by GEROS as exceptional and from other sources. The function is capable of producing high-quality evaluations that are useful and credible.
- The limitations of the evaluation function centre around the wider enabling environment and ensuring accountability and monitoring for roles and responsibilities, including funding and reporting lines, which can limit consistent achievement of independence, credibility and utility. Further risks concern the limitations to the management use of evaluation evidence in planning and implementation of programming, and the interactions between the evaluation function and other knowledge functions within UNICEF.
- The EO has responsibilities for addressing some of the limitations but also needs to be empowered and supported by other elements within the enabling environment, including UNICEF leadership and management in other functions, for its efforts to have maximal effect.

With input from the self-assessment of the EO and using the evidence presented in this report, we summarize progress towards achieving the specific aims of the *2018 Evaluation Policy* and categorized these into mostly or entirely achieved; in progress; and little or no progress. Only five areas have seen little or no progress. This reflects the evidence presented in this report that UNICEF has responded to the *2018 Evaluation Policy* with a seriousness of purpose and to good effect.

Independence
<p>Mostly or entirely achieved</p> <ul style="list-style-type: none"> • The aim to ensure that the Director of Evaluation and Deputy Director have substantial evaluation experience has been achieved • All evaluation Management responses are now on UNICEF's website
<p>In progress</p> <ul style="list-style-type: none"> • A dual reporting line for Regional Evaluation Advisers has been established (although the matrix management model does not unquestionably enhance independence of the function) • The Theory of Change has been developed but not used as a strategic tool to drive the function forward
<p>Little or no progress</p> <ul style="list-style-type: none"> • The recommendation to develop adequate impartiality provisions and safeguards for the behavioural independence of staff with responsibility in planning, commissioning and managing evaluations has not been delivered

<ul style="list-style-type: none"> • The Global Evaluation Committee has not, to our knowledge, been revamped as a platform for substantive collaboration between the EO and UNICEF Management towards collectively ensuring the culture of evaluation • A policy implementation strategy was developed but not used
Governance/management
<p>Mostly or entirely achieved</p> <ul style="list-style-type: none"> • Key Performance Indicators have been established and used on evaluation coverage at country level and on the budget allocated • The Global Evaluation Plan, as envisaged, now includes strategic corporate-level evaluations, as well as evaluation syntheses and meta-evaluations
<p>In progress</p> <ul style="list-style-type: none"> • There are Regional Costed Evaluation Plans but practices and usefulness vary
Finance
<p>Mostly or entirely achieved</p> <ul style="list-style-type: none"> • All EO and Regional Evaluation staff are now (as of December 2022) funded through Regular Resources (with less security at country level) • The multi-country evaluation specialist model has been rolled out
<p>In progress</p> <ul style="list-style-type: none"> • The allocation of 1 per cent has focused attention and been associated with increased funding but is fraught in a number of ways
Human resources
<p>Mostly or entirely achieved</p> <ul style="list-style-type: none"> • An EO Deputy Director role has been established • Director and Deputy Directors have substantial evaluation experience • REAs have been established • Teams of full-time evaluation advisers and specialists have been established at regional and country levels • Provision has been made for the rotation of evaluation staff to support career progression and behavioural independence
<p>In progress</p> <ul style="list-style-type: none"> • The role of Director of Evaluation and its relationship with the Executive Director and the Executive Board have been formally clarified. Given importance, progress should be sustained
Quality
<p>Mostly or entirely achieved</p> <ul style="list-style-type: none"> • An external evaluation advisory committee has been established • Standard guidance for quality assurance has been developed • A real-time Quality Assurance mechanism is in place under the management of the Regional Evaluation Adviser • The time frames for the implementation and closure of recommendations that address strategic and corporate-wide issues have been revised
<p>In progress</p> <ul style="list-style-type: none"> • There has been limited progress towards revising criteria that differentiate evaluations from other products (such as researcher reviews) and that consider a more consistent application of the taxonomy in the titles of evaluation reports
<p>Little or no progress</p> <ul style="list-style-type: none"> • The aim of developing evaluation guidelines from evaluation inception to completion has not been achieved • Little progress has been made towards developing quality standards for evaluation recommendations
Management
<p>Mostly or entirely achieved</p> <ul style="list-style-type: none"> • EO evaluation advisers and specialists are rotated to enable career progression and diversity of perspectives

- Highly reputed specialists in the subject matter of the evaluation have been included in appropriate evaluation teams responsible for carrying out evaluations
- Criteria and Key Performance Indicators identified in the *Evaluation Policy* are adequately monitored and reported upon

Recommendations

The recommendations listed here reflect the recognition that UNICEF is continuing to make good progress towards strengthening its evaluation function to support UNICEF's mandate to advocate for the protection of children's rights, to help meet their basic needs and to expand their opportunities to reach their full potential. Therefore, many recommendations seek to continue or amplify existing approaches. There are other areas where new challenges and new opportunities are apparent that may require a change of direction.

Recommendation 1

The main aim of the 2018 *Evaluation Policy* to support independence, credibility and usefulness through centralized and decentralized evaluation structures should be continued. There is room for further strengthening of roles, responsibilities and accountabilities, particularly at the decentralized level.

Specific complementary measures that should be considered in support of Recommendation 1:

1. Evaluation accountability, reporting and monitoring arrangements should be clarified to help the Evaluation Office to support the aims of the *Evaluation Policy*, particularly at regional and country levels. Clarity should be achieved on how the Evaluation Office can help to ensure that regional directors and country representatives support and monitor the delivery of their accountabilities under the *Evaluation Policy* and how the Evaluation Pooled Fund is governed, allocated and distributed at the regional level (see Recommendation 2 for more detail).
2. Revisit and use an evaluation Theory of Change as a live document. In the context of updating the new evaluation policy, it would be a good moment to make clear what the Evaluation Office and regional and country-level staff can and should do and where they need enabling environment support from UNICEF management.
3. To ensure that the overall portfolio of evaluations aligns and supports the mission and strategic objectives of UNICEF, and based on systematic assessment of gaps in evidence, the Evaluation Office should take stock of where evaluations of existing or planned programming may help to fill operationally important evidence gaps. This should be used to inform the planning of priority topics for evaluation.³⁵ UNICEF's Office of Research – Innocenti – could also play a useful role in helping to ensure that priorities are optimized.
4. Efforts at the global, regional and country levels to strengthen a culture of evaluation within the organization – reflected in words and actions as well as inscribed in policy and protocol – should be sustained and strengthened. To do so, the Evaluation Office could draw on the use of Communities of Practice as conducted in the World Food Programme and practices to strengthen learning from evaluation at USAID (Annex 4).
5. The Evaluation Office should identify ways to strengthen the accessibility and usefulness of evaluation products, including producing short summaries, infographics and presentations

³⁵ Potentially useful resources for this include the Cochrane handbook for systematic reviews of interventions (<https://training.cochrane.org/handbook>), which provides guidance on identifying research gaps, and the James Lind Alliance (<https://www.jla.nihr.ac.uk/>), which provides a road map for prioritizing evaluation topics. Lessons from UNDP, which uses an 'M E Sandbox' to support flexibility and learning, could also be considered.

(as well as longer and more-detailed reports). These could be either developed in-house, developed in collaboration with the communications function or required as standard deliverables in terms of reference for decentralized evaluations.

6. The routine monitoring of evaluation quality (GEROS) and management responses are important ways to direct attention to quality and use and should be maintained (while recognizing that they are not intended to be a complete source of information about quality and uptake).

Recommendation 2

UNICEF management should ensure the sufficient and predictable funding of the evaluation function. The 1 per cent target should be clarified, including how it is calculated and what is included. Governance and criteria for allocation and use of the Evaluation Pooled Fund should be clarified.

Specific complementary measures that should be considered in support of Recommendation 2:

1. The 1 percent should be clarified, including and how it is calculated and what is included. Moreover, it should be used as a mechanism to ensure that sufficient and predictable funding is made available by management for evaluation in a timely manner. To ensure real and perceived independence, the evaluation function should not be responsible for fundraising to fulfil its evaluation plans (fundraising may be appropriate, however, for national capacity development activities, for example).
2. The Evaluation Office should build greater clarity about the optimal level of more sustainable funding for the evaluation function (including human resources for evaluation). This should be raised to the Executive Board for their consideration to allow them to arrive at a considered judgement about whether UNICEF is investing sufficiently in evaluation.
3. UNICEF management should ensure that the funding modalities result in a more predictable flow of funds that is better aligned with programming cycles to support the more efficient and effective use of evaluation funds.
4. A set of actions to improve resource allocation and monitoring through the use of KPIs and reporting should be considered. Areas for potential consideration could include: Evaluation Office to continue to produce quarterly KPI Corporate Reports; regional offices to produce quarterly disaggregated KPI Regional Reports and share and discuss these periodically with representatives; Costed Evaluation Plans to be used to provide greater visibility and accountability as soon as practical by: the Evaluation Office uploading CEPs on UNICEF global database and by regional offices and the Director of Evaluation periodically providing the Executive Director with a summary of status of CEP implementation; regional offices to provide the Evaluation Office with analytical inputs on use of evaluations in CPDs for EO to synthesize for the Executive Director Annual Report.

Recommendation 3

The evaluation function should build on its added value by identifying, together with other knowledge functions in the organization, specific and focused ways to work more closely together.

Specific complementary measures that should be considered in support of Recommendation 3:

1. The Evaluation Office has established a commitment to initiate a series of one-to-one engagements with the leadership of other knowledge functions, with a view to identifying specific and concrete ways to create synergies across the functions. Specifically, other relevant knowledge functions include: Research, Innovation, DAPM, Internal and External Audit, Office of Emergency Programmes, ICT and Advocacy. The peer review commends and encourages this initiative.
2. UNICEF should retain and strengthen its focus on participatory approaches that involve children and young people, as appropriate and in line with ethical protocols, not only as sources of data, but as key contributors to evaluation planning, design, analysis and production of findings and recommendations.
3. The digital issue will be critical for evaluation in the coming five years and therefore the Evaluation Director should work closely with the ICT Board (soon to become the Digital Transformation Board); conduct annual workshopping around the digitalization of UNICEF and its implications of evaluation; and engage regularly with the Chief Data Scientist to identify implications for data-mining and other opportunities and risks.
4. The Director of Evaluation, working with Heads of Offices, should invest in innovative technologies to advance the use of evaluations, including platforms for connecting evaluation products and evaluators with users of evaluation evidence.
5. Over a period of years, there should be a concordat across the knowledge functions identifying how benefits for the organization, and ultimately for children, were achieved through better collaboration.

Recommendation 4

UNICEF should maintain and further strengthen the improvements that have been achieved in its evaluation of human resources and leadership capacities.

Specific complementary measures:

1. The Evaluation Director should continue to report directly to the UNICEF Executive Director through, *inter alia*, regular one-on-one meetings. The Evaluation Director should continue to have direct access to the Executive Board.
2. The Global Evaluation Committee should promote an enabling environment for evaluation in UNICEF.
3. The practice of appointing an Evaluation Director and Deputy Director with extensive evaluation experience should be continued.
4. The practice of establishing a professional evaluation presence in regional offices should be continued. The Evaluation Office should consider whether the reporting lines and responsibilities of Regional Evaluation Specialists support effective evaluation governance.
5. Building career progression and independence for evaluation practitioners in country offices should evolve alongside the practice of rotating staff members.
6. Action to support staff development and strengthening the practices of commissioning and managing externally provided evaluations should be continued, learned from and, where appropriate, scaled up.

Recommendation 5

UNICEF should use its unique mandate and visibility to work with and influence the external evaluation environment of UN agencies, academics, voluntary evaluation organizations and foundations to promote national evaluation capacities and to mobilize evaluative action in support of UNICEF's mission.

Specific complementary measures:

1. Continue to promote evaluation findings and insights through peer-reviewed journals and conferences.
2. Engage with evaluation societies to promote the findings, insights and aims of UNICEF's evaluation function.
3. Engage corporate and government providers to leverage support for shared evaluation interests.