

Transforming Programming: Evaluation of the implementation of the UNICEF Gender Policy and Gender Action Plans (GAP 2 and GAP 3)

Inception Report

Submitted to UNICEF Evaluation Office
18 April 2024



Commissioned by the UNICEF Evaluation Office, this evaluation is undertaken by Social Development Direct and is expected to occur from November 2023 to September 2024. The Evaluation Manager is Erica Mattellone, and the consulting team undertaking the evaluation includes:

- Celine Mazars
- Alix Clark
- Florianne Gaillardin
- Sonal Zaveri
- Lucia Soldà
- Hanna Smit
- Jo Feather
- Isadora Brizolara
- And National Consultants.

This inception report was submitted to the Evaluation Office on 20 March 2024 and finalized on 1 April 2024.

CONTENTS

1. INTRODUCTION.....	1
1.1 Purpose of the Inception Report.....	1
1.2 Activities conducted at the inception phase.....	1
2. CONTEXT AND OBJECT OF THE EVALUATION	2
2.1 Policy and programming context.....	2
2.2 UNICEF GAP 2 and 3 and UNICEF Gender Policy.....	3
3. PURPOSE, OBJECTIVE AND SCOPE OF THE EVALUATION	4
3.1 The purpose and objectives of the evaluation.....	4
3.2 Key issues emerging from the inception phase	5
3.3 The scope of the evaluation.....	6
3.4 The intended users and use	6
4. APPROACH.....	7
4.1 Analytical approach: Theory-based and appreciative enquiry	7
4.2 Participatory approach.....	8
4.3 Evaluation criteria	9
4.4 Evaluation principles.....	10
4.5 Ethical approach and safeguarding.....	11
5. EVALUATION FRAMEWORK	11
5.1 Evaluation questions.....	11
5.2 Evaluation Theory of Change.....	15
5.3 Evaluation matrix	18
6. METHODOLOGY	18
6.1 Summary of overall methodology and documents produced.....	18
6.2 Analysis of global data and evidence	19
6.2.1 Desk review (documents and data).....	19
6.2.2 Primary data collection	21
6.3 Comparative analysis study.....	22

<i>6.3 Case Study Methodology</i>	25
6.3.1 Case study selection	25
6.3.2 Objectives of the case studies	28
6.3.3 Case Study Preparation	28
6.3.4 Case study methodology	28
6.3.5 Top-up case study	31
6.3.6 Case study analysis and reporting	32
<i>6.4 Data collected in an additional set of country offices</i>	32
<i>6.5 Sampling</i>	33
<i>6.6 Analysis and synthesis for the evaluation report</i>	34
7. DELIVERABLES	35
8. ASSUMPTIONS AND RISKS	37
9. EVALUATION GOVERNANCE, MANAGEMENT ARRANGEMENTS AND EVALUATION TEAM	40
10. WORKPLAN	43

LIST OF FIGURES

Figure 1: GAP theory of change	4
Figure 2: Evaluation theory of change	16
Figure 3: Evaluation theory of change assumptions	17
Figure 4: Evaluation methodology phases	19
Figure 5: Full Case study methodology	29
Figure 6: Summarized GESI Continuum	35
Figure 7: Team diagram.....	40

LIST OF TABLES

Table 1: Users of the evaluation	7
Table 2: YAG involvement	9
Table 3: DAC criteria applied to the evaluation	9
Table 4: Proposed amendments to the evaluation questions in the ToRs.....	13
Table 5: Provisional list of KIIs/FGDs (to be finalized with the Evaluation Office).....	21
Table 6: Proposed organizations for comparative analysis in relation to criteria.....	23
Table 7: Proposed KIIs/FGDs for comparative analysis.....	24
Table 8: Proposed case study countries	27
Table 9: Proposed stakeholders for in-country KIIs	29
Table 10: Proposed stakeholders for in-country FGDs	31
Table 11: Proposed participants for in-country participatory workshops	31
Table 12: Proposed KIIs for in-country supplementary data collection.....	33
Table 13: Overall number of participants to KIIs, FGDs and workshops	33
Table 14: Key deliverables.....	36
Table 15: Risks and mitigation measures.....	37
Table 16: Assumptions and mitigation measures.....	39

TABLE OF CONTENTS

AAG	Adolescent Advisory Group
ADAP	Adolescent Development And Participation
AP	Adolescent Panel
CCC	Core Commitments for Children
CEDAW	Convention on the Elimination of All Forms of Discrimination against Women
CO	Country Office
CRPD	Convention on the Rights of Persons with Disabilities
CRC	Convention on the Rights of the Child
CSO	Civil Society Organization
DAC	Development Assistance Committee
DAPM	Division of Data, Analytics, Planning and Monitoring
DTL	Deputy Team Lead
ERG	Evaluation Reference Group
EAPRO	East Asia and the Pacific Region Office
ECARO	Europe and Central Asia Regional Office
EMOPS	Office of Emergency Programmes
ESARO	East and Southern Africa Regional Office
EO	Evaluation Office
EQ	Evaluation Questions
ERG	Evaluation Reference Group
FCAS	Fragile and Conflict-Affected settings
FGD	Focus Group Discussion
GAP	Gender Action Plan
GBViE	Gender-Based Violence in Emergencies
GESI	Gender Equality and Social Inclusion
GEM	Gender Equality Marker
GEROS	Global Evaluation Reports Oversight System
GII	Gender Inequality Index
GRP	Gender Programmatic Review
GTP	Gender Transformative Programming
HQ	Headquarters
IRB	Institutional Review Board (ethics committee)
KII	Key Informant Interview
LACRO	Latin America and the Caribbean Regional Office
LMIC	Low- and middle-income countries
OECD DAC	Organization for Economic Co-operation and Development's Development Assistance Committee
MOPAN	Multilateral Organizational Performance Assessment
MENARO	Middle East and North Africa Regional Office,
NGO	Non-Governmental Organization
NLP	Natural Language Processing
QA	Quality Assurance
ROSA	Regional Office for South Asia
RAGs	Regional Gender Advisors
RO	Regional Office
SDDirect	Social Development Direct
SDGs	Sustainable Development Goals
SEAH	Sexual Exploitation, Abuse and Harassment
TL	Team Lead
ToC	Theory of change
ToR	Terms of Reference
UNEG	United Nations Evaluation Group
UN-SWAP	UN System-wide Action Plan on Gender Equality and the Empowerment of Women
UNEG	UN Evaluation Group
WCARO	West and Central Africa Regional Office

WHO	World Health Organization
YAG	Youth Advisory Group

1. INTRODUCTION

1.1 Purpose of the Inception Report

The UNICEF Evaluation Office has contracted Social Development Direct (SDDirect) to evaluate and assess the effectiveness of the implementation and the results achieved through the new Gender Policy (2021-2030) and the Gender Action Plans (GAPs) 2018-2021 and 2022-2025. This evaluation is part of the Plan for Global Evaluations attached to the Strategic Plan, and the evaluation report will be presented to the UNICEF Executive Board in 2025. The scope of the evaluation is global, and the Evaluation Office will ensure the independence, impartiality and credibility of the evaluation process as per UNICEF Evaluation Policy (2023).¹

This inception report outlines our understanding of the evaluation and presents how we will conduct it. It ensures that the SDDirect team, the UNICEF independent Evaluation Office, the Gender Team, and the Evaluation Reference and Advisory Groups members share the same understanding of the objectives, scope, evaluation questions and methodology, risks and mitigation measures, and the theory of change. It also serves as a guiding framework for data collection and analysis.

1.2 Activities conducted at the inception phase

The evaluation started in November 2023, with the inception phase from December 2023 to February 2024. This phase enabled the team to engage with the evaluation questions and existing evidence on the GAPs and identify critical issues that must be considered throughout the evaluation, presented in section 3.2.

The following activities were conducted:

- Preliminary review of UNICEF documentation, including the current UNICEF Strategic Plan (2022-2025), previous Strategic Plans (2018-2021, 2014-2017), GAPs documents and reports, Gender Policy documents, past evaluations and management responses (including the first evaluation of the first Gender Policy), UNICEF activity reports, amongst others. Overall, 50 priority documents were reviewed.
- Review the existing monitoring system and available data and analyse them to understand better what data exists and how they could be analysed to respond to the evaluation questions.
- The Constitutions of the Evaluation Reference Group and the Youth Advisory Group, as well as meetings to introduce the evaluation, its objectives, methodology, and team, provide an overview of the groups' roles, and seek their input and suggestions.
- In-depth engagement with UNICEF staff at various levels to support the identification of key themes to explore in the evaluation and ensure its utilization focus:
 - Weekly management and technical meetings were held with the Evaluation Office, Gender team and focal points for the evaluation.
 - 15 key informant interviews were conducted with UNICEF staff at all levels, including the HQ Gender team staff, the regional gender advisors (RGAs) and two staff from country offices. These semi-structured interviews identified expectations for this evaluation, the achievements and remaining challenges since the last evaluation and the perceived primary shifts between GAP 2 and 3; responses to the questions were then analysed and thematically coded using the Dedoose software. A list of interviewees is included in Annex 1: Stakeholders consulted in the inception phase.
- Based on the interviews, documents, and data analysed, the evaluation team developed a draft theory of change (ToC) to guide the evaluation and provide a basis for discussion.
- The ToC was amended and developed during a series of participatory workshops with

¹ See UNICEF (2023). [Revised evaluation policy of UNICEF.](#)

various groups:

- Two online workshops with regional gender advisors were held in January 2024.
- A series of mini-workshops were held in New York over four days to finalize the theory of change and underlying assumptions.²
- Debrief meetings were held to present emerging critical themes and the final ToC.

There has been a high level of engagement with the UNICEF team during this evaluation phase. More than 20 people attended the final ToC workshop in New York, and further inputs were received by email and offers to support the process moving forward.

2. CONTEXT AND OBJECT OF THE EVALUATION

2.1 Policy and programming context

UNICEF's commitments to gender equality

UNICEF's efforts to promote gender equality are driven by the 1989 Convention on the Rights of the Child (CRC), the 1979 Convention on the Elimination of All Forms of Discrimination against Women (CEDAW) and the 2006 Convention on the Rights of Persons with Disabilities (CRPD), as well as the 1986 Declaration on the Right to Development and the 1993 Declaration on the Elimination of Violence against Women. In emergencies, UNICEF's work is informed by the Core Commitments for Children in Humanitarian Action (revised in 2020), which promotes gender equality in all humanitarian actions.

As a participating entity of the UN System-wide Action Plan (UN-SWAP) on gender equality and women's empowerment, UNICEF annually reports on the 17 common performance indicators. UNICEF contributes to the implementation of the Sustainable Development Goals (SDGs), in particular, SDG 5, to achieve gender equality and empowerment for all women and girls, with a focus on target 5.1 on ending all forms of discrimination against women and girls everywhere and 5.2 on ending all forms of violence against and exploitation of women and girls.

Gender equality became more visible in the organization. It was explicitly integrated as a cross-cutting issue in the UNICEF Strategic Plan (2006-2009) and the UNICEF Policy on Gender Equality and the Empowerment of Girls and Women (2010).³

The first Gender Action Plan (2014-2017) aligned with the 2014-2017 Strategic Plan programmes. It focused on mainstreaming gender in programmes, results and performance monitoring, capacity and systems strengthening, resources and partnerships, and recommended interconnected programming to improve the well-being of adolescent girls.

Following the principal recommendation of the evaluation of GAPs 1 and 2 in 2019⁴, which urged UNICEF to “significantly upscale the ambition and aspiration of its vision for gender equality, commensurate with its mandate and status as the world's defender of child rights”, UNICEF developed the 2021-2030 Gender Policy,⁵ which “commits UNICEF to a bolder and more ambitious vision for gender equality and the empowerment of all children, adolescents, and women.” It explicitly articulates the organization's gender transformative ambitions. It compels UNICEF to “work actively to remove the underlying structural barriers – such as harmful social norms and gendered power systems – that perpetuate inequalities”. It acknowledges gender diversity and

² Participants included representatives of the following teams: Gender, Programme Group, DAPM, HR, Culture & Diversity, EMOPS, and ADAP; representatives of the staff group Gender Push were also interviewed.

³ See UNICEF (2005). [The UNICEF medium-term strategic plan, 2006-2009 Investing in children: the UNICEF contribution to poverty reduction and the Millennium Summit agenda](#); UNICEF (2010). [Working for an Equal Future UNICEF Policy on Gender Equality and the Empowerment of Girls and Women](#).

⁴ See UNICEF Evaluation Office (2019). Realizing Potential: Evaluation of UNICEF's Gender Action Plans.

⁵ UNICEF (2021). [UNICEF Gender Policy 2021-2030](#).

proposes to uplift adolescent girls' rights and deepen partnerships with women and girls' rights organizations by 2030.

Gender equality is a key theme in UNICEF's work: It is a cross-cutting focus, a core principle, and a change strategy (gender transformative programming) within the 2022–2025 Strategic Plan.

Elements of the context that impact work to promote gender equality

Since the development of GAP 2 in 2019, it has become apparent that the situation of women, girls and children is not improving as fast as envisaged in the SDG targets, and SDG 5 is off-track⁶. The COVID-19 pandemic has affected women and girls disproportionately. At the same time, the world witnesses a backlash against women's rights, with an increasing number of countries introducing gender discriminatory laws, which are making it increasingly difficult to engage in transformative gender work – a challenge that many international organizations are grappling with.

In parallel, increased evidence on how to tackle gender norms and a growing consensus on the need to adopt gender transformative approaches in programming and work with women, girls, and youth-led organizations to ensure sustainable structural changes are facilitating the work.

In this context, the United Nations has reaffirmed the need to accelerate efforts to achieve gender equality through the Decade of Action towards the 2030 Agenda. The UN Reform accompanies this new impetus.

2.2 UNICEF GAP 2 and 3 and UNICEF Gender Policy

The GAPs 2 and 3 adopt a whole-of-institution approach, spanning programmes, institutional structures, programmes and systems.

The vision for GAP 2 (2018-2021) was to accelerate gender programming and take it to scale. It also aimed to intensify gender integration in institutional structures and systems by expanding gender capacity and expertise across UNICEF at all levels, reinforcing the importance of gender analysis, data, research and evidence. It also had a particular focus on programmatic priorities for adolescent girls in the areas of health, secondary education, child marriage, gender-based violence in emergencies (GBVIE), and menstrual health and hygiene.

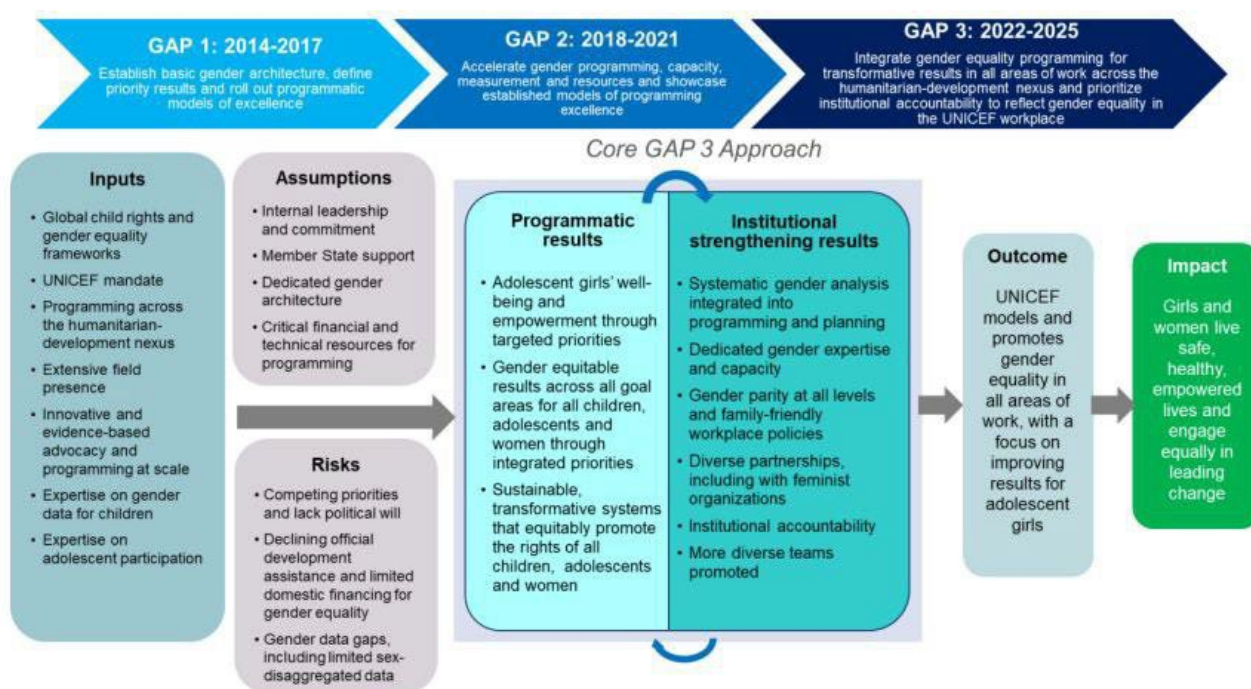
GAP 3 (2022-2025) builds on the UNICEF Gender Policy 2021-2030 and seeks to address the challenges faced in implementing GAP 2 in response to the recommendations made in the 2019 evaluation. It sets out to do this by focusing more on institutional accountability and ownership in all contexts in which UNICEF operates, addressing underlying structural barriers to change power dynamics and social norms, and promoting adolescent girls' leadership and well-being. It also focuses on the organization's work on advocacy, innovation, and partnerships with feminist organizations and networks, as well as data, research and analysis. GAP 3 was developed in a participatory way and involved extensive consultations across the organization and externally to "ground truth" the documents and ensure ownership and synchronicity across UNICEF's decentralized structures".⁷

The GAP is guided by a theory of change (ToC) grounded in a twin-track approach, tackling both programmatic and institutional components.

⁶ UNICEF Division of Data, Analytics, Planning and Monitoring (2023). [Progress on Children's Well-Being: Centering child rights in the 2030 agenda – For every child, a sustainable future](#).

⁷ UNICEF Gender Unit, January 2024, Preparing for an independent evaluation of UNICEF's Gender Policy and Action Plans 2018 to the present: A background paper.

Figure 1: GAP theory of change



The GAP ToC is guided by programmatic priorities that support the integration of gender equality in all areas of work across the humanitarian-development nexus, focusing on improving results for adolescent girls. The ToC also reflects the integration of gender equality in organizational processes and systems guided by the institutional enablers framework. It is accompanied by an indicator matrix, which is reported to the board as part of the Strategic Plan reporting processes

3. PURPOSE, OBJECTIVE AND SCOPE OF THE EVALUATION

3.1 The purpose and objectives of the evaluation

This evaluation aims to assess UNICEF's performance and results in implementing GAP 2 and GAP 3, the new Gender Policy, and the humanitarian, development, and peace nexus from 2019-2023. It is conceived as a learning and accountability exercise and is part of the Plan for Global Evaluations linked to the Strategic Plan. The evaluation report will be presented to the UNICEF Executive Board in 2025.

The specific objectives of the evaluation are to:

- To assess the relevance and coherence of the current Gender Policy and GAP3 in supporting gender equality and adolescent girls' empowerment from the point of view of UNICEF's mandate (i.e., CRC, CEWAD, CRPD, CCCs) and the organization's commitments to the UN-SWAP, as captured in the Strategic Plan in development and humanitarian settings.
- To assess the coherence, effectiveness, efficiency, and sustainability of implementing gender-related organizational changes and performance enablers of the GAPs across UNICEF policies, practices, systems, programming processes, organizational commitments, and accountability mechanisms.
- To determine the extent to which UNICEF meets the programmatic results set for all children, adolescents and women on gender equality across the five Goal Areas of the Strategic Plan and in both development and humanitarian settings.

The recommendations from the evaluation will:

- Help to reimagine the current GAP framework and ToC and inform its implementation.
- Influence UNICEF’s strategic direction on gender equality, including informing the development of the next Strategic Plan and a possible new GAP or alternative.
- Help to strengthen organizational arrangements to create an enabling environment for gender equality, including assessing equal representation and capacity.
- Allow UNICEF to assess the results (positive or negative) achieved with partners for gender-targeted interventions, including outlining enabling and hindering factors.
- Generate learning and address internal accountability of UNICEF’s progress towards gender equality.

3.2 Key issues emerging from the inception phase

Key informant interviews at the inception phase and a series of workshops conducted with HQ and regional staff allowed the team to discuss the ToC that guides this evaluation and, importantly, identify key issues that informed the development of assumptions in the ToC. The problems identified fall primarily under two themes: the gender transformative agenda and the challenges related to the institutional pathway, which are presented below.

The gender transformative agenda:

Despite an acknowledgement that staff buy-in on the gender mandate is increasing, aided by the Gender Policy and GAP, there were different perceptions on the extent to which leadership at all levels fully endorses and champions gender as part of UNICEF’s mandate, focusing on child rights.

In addition, our initial engagement with UNICEF staff at various levels suggest there are different understandings and levels of awareness within the organization of what constitutes the gender transformative agenda (something that was observed in 2019 during the last GAPs evaluation). The term can create confusion and was said to cause some resistance among some staff members, government partners and UNICEF’s board members. As a result, the GAP, which had to be presented to the Executive Board, uses adapted language, such as “gender equality results that are transformative”. The Gender Policy, on the other hand, which is ambitious in language and overtly gender transformative, is not consistently referred to within the organization. In parallel, there are multiple guiding documents and tools promoting gender transformation in programmes, and some donors request the programmes they fund to include gender transformative indicators that go beyond those used as standard by UNICEF. The evaluation will explore this tension in greater depth.

Ongoing challenges related to the institutional change pathway:

Initial data collected and analysed during the inception phase suggests that UNICEF has made good progress on implementing the institutional enablers (as defined in GAP 3 and in the inputs section of our theory of change⁸) over the period covered by the evaluation. For example, between 2019 and 2021, the number of UNICEF offices meeting the GAP institutional standard increased from 63 to 98, the M&E system was improved, as well as improvements in capacities to integrate gender were noted. However, according to the new GAP Standard, only 12% of country offices meet the threshold related to the GAP institutional enablers elements. Remaining challenges hampering the operationalization of the GAP flagged by the annual GAP reports and consultations with staff conducted at inception phase include: (a) insufficient resourcing of dedicated gender staff at regional and country levels as well as capacity gaps; (b) data completeness and accuracy; (c) limited ability of the current indicators to capture system-level changes on gender beyond process implementation, as well as changes in social norms and adolescent girls’ outcomes; (d) the perception that the M&E system is over-complicated and time-consuming; and (e) mixed perceptions of the extent to which the GAP is able to influence sectoral and country plans.

⁸ See section 5.2 below.

In addition, workplace discrimination was highlighted in an independent task force report (2020) that issued internal recommendations on the topic, including salary-based discrimination. Interviews conducted during the inception phase revealed workplace issues that questioned the extent to which UNICEF is modelling the gender transformation it calls for in its policies and practices within the office, beyond the attention to gender parity in the workforce and addressing Sexual Exploitation, Abuse and Harassment (SEAH) issues. Cases of gender bias, limited attention to women’s well-being in the workplace and resistance to change were quoted. Finally, issues of limited prioritization and focus on behalf of some senior management at various levels of the organization were mentioned, echoing the challenges related to accountability mentioned in the previous evaluation.

This said, in contrast to the 2019 evaluation of GAP 2, which faced a paucity of data on gender, our team identified a greater volume of quantitative data than anticipated, which reflects an increased focus and effort on the part of UNICEF to capture gender-related work better and will constitute a valuable source of information for the evaluation.

A synthesis of key learnings from the inception period and their implications for the way we interpret and respond to the evaluation questions is presented in the Evaluation Matrix (see Annex 4: Evaluation Matrix).

3.3 The scope of the evaluation

This evaluation will be global in scope, with in-depth case studies in a sample of countries (see section 6.3 on the case study methodology) and spans from 2018 to 2023, covering the implementation of GAP 2 and half of GAP 3.

The evaluation will respond to most of the questions identified in the original terms of reference (ToR), as they are directly linked to the recommendations made during the previous evaluations, but the evaluation team proposes a few amendments and consolidation of some questions, which are presented in section 5.1 (Evaluation questions).

The evaluation will cover three main components:

- **Gender Policy and GAP conceptual framework:** The evaluation will examine the strategic positioning of gender in UNICEF’s new Strategic Plan, analysing the Gender Policy, GAP conceptual framework and ToC.
- **Institutional enablers and their implementation:** The evaluation will assess how the GAP is implemented, including programming processes, monitoring systems, financial resources, leadership and accountability, at the headquarters, regional and country levels.
- **Programmatic results and integration of gender equality across programmes:** The evaluation will examine overall programmatic results based on available indicators and implemented activities, and assess the outputs and the contribution to outcomes as defined in the GAPs in the selected case studies. In particular, the evaluation will assess to what extent UNICEF has mainstreamed gender equality across its five Goal Areas in the Strategic Plan.

Our understanding of the scope of each evaluation question is presented in the Evaluation Matrix, in section 5.3 below, before presenting the various assumptions embedded in the ToC.

3.4 The intended users and use

As per the ToR, the primary users of this evaluation are “UNICEF employees at all levels, government and partners working with UNICEF to advance the rights of women and girls, women and girls-led organizations, and young people”.

The table below presents the primary and secondary users of this evaluation as well as the type

of use.

Table 1: Users of the evaluation

Primary audience of the evaluation report	Type of use
UNICEF Gender Section in the Programme Group	To inform possible adaptations to the current GAP implantation modalities, as well as analysis and recommendations to support the development of future gender frameworks.
UNICEF senior leadership and the board	To take stock of the achievements of the GAP and define future strategies as well as hold UNICEF accountable for its commitments. The report will be presented to the UNICEF Executive Board in February 2025.
Country Offices	To strengthen implementation of policies and programmes and provide strategic guidance to the next Country Programmes.
Other UNICEF staff at all levels	To strengthen gender integration in their work.
Secondary users	Type of use
Government and partners working with UNICEF to advance the rights of women and girls, women and girls-led organizations, and young people, other UN agencies	Shared learnings on how to integrate gender equality in programmes and organizations, and on building a culture of accountability.

4. APPROACH

4.1 Analytical approach: Theory-based and appreciative enquiry

Our approach is theory-based⁹ and we will use an analytical framework described in the theory of change (ToC) developed during the inception phase¹⁰, which helped us to refine the change pathways that were implicit in the existing GAP ToC; define explicit causal assumptions linking GAP support to identifiable results at the output, outcome and goal levels; and refine the evaluation questions to be investigated, against which findings and conclusions will be reported.

The ToC and assumptions developed have informed the development of the evaluation matrix, including sources of information and data collection tools. It reflects a shared understanding among key UNICEF stakeholders and the evaluation team of the intended causal chains underpinning the UNICEF Gender Policy and Gender Action Plans. It provides the theoretical framework for applying contribution analysis to assess causal linkages and infer the contribution UNICEF has made to the observed outputs and outcomes.

Appreciative inquiry is an approach to organizational culture change, that focuses primarily on strengths. We will adapt this approach to evaluate the GAPs (2 and 3) and Gender Policy, identifying lessons on what is working well which will inform recommendations to accelerate organizational change. The approach includes the following phases:

- **Discovery:** Under the discovery phase the approach involves the exploration of processes that work well. The consultation, review and data collection that took place during inception, have allowed us to identify country offices that are performing well on integrating gender across their programmes and operations based on the UNICEF GAP standard index data. This will allow us to draw out processes, and approaches that are working well through primary data collection.

⁹ Drawing on a methodology described in: Mayne, John, 'Contribution Analysis. An approach to exploring cause and effect', ILAC Brief 16, 2008.; Mayne, John, 'Revisiting Contribution Analysis', Canadian Journal of Programme Evaluation, December 2019.

¹⁰ The process to develop the ToC is described in section 1.2

- **Dream:** Under the dream phase the approach aims to explore the aims and the vision and what is needed to achieve the vision. As this evaluation aims to assess the implementation of the GAPs and Gender Policy, in which a ‘dream’ is already articulated, we have adapted this phase. During the inception phase a participatory ToC process provided space for the articulation of the vision and causal change pathways which the evaluation will assess. We will consider whether the vision as outlined in the GAPs and Gender Policy is shared among staff, is aligned to other framework visions, is relevant to the needs and rights of women and children, and whether anything is missing or irrelevant, providing participants space to dream.
- **Design:** Under this design phase, the approach calls for identifying and co-designing processes that could work well to achieve the vision identified above. These processes have been defined as performance enablers under the GAP 3 and this evaluation will assesses the implementation of these performance enablers outlined in the GAP to answer Evaluation Question 2. The evaluation will also provide space to co-create recommendations on implementing the performance enablers moving forward.
- **Delivery:** Under this phase the approach focuses on implementation of the vision and how to sustain. Following the data collection phase, the evaluation will draw out lessons on what is working and what has been achieved and make recommendations for how to sustain and accelerate progress.

4.2 Participatory approach

Working with the Evaluation Office

The evaluation team has worked closely with the UNICEF Evaluation Office (EO) throughout the inception phase to ensure impartiality, independence and credibility of the work undertaken. The Evaluation Manager from the EO has coordinated with the Gender team responsible for coordinating the GAP, Programme Group and other Headquarters Divisions, Regional Offices and Country Offices, who will be responsible for ensuring access to information and key stakeholders to ensure that the evaluation will produce relevant and reliable findings and actionable recommendations. A member of the Evaluation Office has been embedded in the evaluation team to provide access to information and key stakeholders for the inception phase, and another member will be supporting with Natural Language Processing for analysis of global data and evidence (see section 6.2.1).

Working with the Evaluation Reference Group (ERG)

An ERG has been established to ensure ownership of the evaluation and provide expert advice and inputs to the evaluation team at key moments during the delivery of the assignment. The group features 18 members, including UNICEF staff and external stakeholders. The main purpose of the group is to support quality assurance during the evaluation process, methodology and key deliverables during the evaluation process, and support and validate key recommendations in the follow-up phase. The full list of members of the ERG as well as a description of the group’s responsibilities for the evaluation can be found in Annex 13: ERG role and list.

Working with the Youth Advisory Group (YAG)

During the inception period, the EO established a Youth Advisory Group (YAG) to facilitate **active youth engagement** at all stages of the evaluation process. The aim of this group is to ensure the evaluation addresses the issues that adolescent girls and boys prioritize that adults may not be aware of, and to ensure their voices are heard and their experiences reflected. This will increase the evaluation’s utility and relevance.

The YAG is made up of 12 members: nine young women and three young men aged between 18 and 25, representing 10 countries (India, Syria, Zimbabwe, Greece, France, Cameroon, Barbados, Canada, Morocco, United States). Efforts have been made to ensure a diversity of young women and men are included in the YAG, through for example including representation from a variety of UNICEF programmes, countries and regions and different age ranges. The evaluation Deputy

Team Lead will work in close collaboration with the UNICEF Evaluation Office to ensure meaningful engagement with this advisory group from design stage through to validation and dissemination. More information about the members of the YAG and the group’s responsibilities can be found in Annex 14: YAG role and list. The YAG will advise the evaluation team at key points throughout the evaluation. For example, they will have the opportunity to feedback on key deliverables and to inform the methodology. The group will be engaged in person at four key stages of the evaluation:

Table 2: YAG involvement

Engagement	Purpose and Inputs
Kick off Call	An initial orientation call with the YAG was hosted during inception to introduce the evaluation, outline roles and establish ways of working.
Inception Report: presentation	A second meeting was held with the YAG to present the methodology for the evaluation in more detail, placing specific targeted focus on the engagement of young people through the case studies. The YAG provided feedback on the sampling for and approach to data collection with young people. This feedback will be considered during the design of the focus group tool for the adolescents. The draft tool will be shared with the YAG for their review and comment ahead of its use in the pilot case study.
Presentation of preliminary findings and co-creation of recommendations	A third meeting will be held to present the preliminary findings from the case studies, with particular focus on the feedback from the young people consulted during the process. The YAG will have the opportunity to reflect on these findings and co-create recommendations for UNICEF.
Presentation of Findings and consultation on youth-friendly products	A fourth meeting will present the overall evaluation findings to the YAG for their reflection. In this meeting the focus will be placed on engaging the YAG to advise on the creation of youth friendly products.

The YAG will not have direct involvement in specific case studies or data collection due to a lack of geographic overlap and limited resources to train and upskill the YAG members. We could explore opportunities for YAG members to shadow or be involved in global interviews where possible.

Another illustration of the participatory nature of our approach is participatory is the involvement of country offices to review our data collection protocol and provide inputs to contextualize it and ensure the evaluation will provide actionable recommendations, in line with our utilization-focus approach.

4.3 Evaluation criteria

The table below describes the criteria applied to the Gender Policy and GAPs evaluation, drawing on DAC criteria as well as UNICEF’s definition of Impact.

Table 3: DAC criteria applied to the evaluation

Relevance	Extent to which the GAP is responsive to the various contexts in which UNICEF operates (development or humanitarian, LMIC or MIC countries and various continents), and to the diverse gender realities and needs; extent to which it is grounded in evidence of what works.
-----------	--

Coherence	Internal coherence will focus on the alignment, synergies and interlinkages between the GAP and other UNICEF normative frameworks, including the coherence between the GAP and the Gender Policy. External coherence will examine the alignment of the GAP with broader UN commitments and to governments.
Effectiveness	The extent to which the results intended by the GAPs 2 and 3 have been realized (as far as feasible to discern at this stage) and whether contribution has been demonstrated towards results which could have been reasonably expected. Analysis of the varying importance of the results and differential results across groups and sectors and understanding the enabling and hampering factors that influence results.
Impact	The evaluation will adopt UNICEF definition of impact presented below, and seek to identify the plausible contribution to impact results: The “positive and negative, direct or indirect, primary and secondary, short, medium or long-term change in the lives of children and families produced by an intervention.” ¹¹ We provide details on the scope of our impact analysis in section 5.1 below. However, it is important to note that the evaluation budget and scope is not sufficient to ascertain the impact of UNICEF programmes on adolescent girls at global level - we will rely on the analysis of secondary data from the GAP indicators.
Sustainability	The robustness of the institutional enablers will give an indication of the sustainability of the changes observed, as well as the strength of UNICEF partnerships with governments and the civil society sector.

More information on our understanding of the criteria and how they will be applied is presented in the evaluation matrix.

4.4 Evaluation principles

The evaluation will be guided by the following principles:

- It is **utilization focused**, so that the assessment findings and process can be used to inform decision making and improve programming with the view to benefit rights holders, in particular adolescent girls and young women, and adolescent boys and young men, and to strengthen the integration of gender within UNICEF structures and systems. The evaluation will also take a forward-looking perspective with a strong focus on lesson learning.
- **Co-creation and co-ownership**: Throughout the evaluation process, the SDDirect team will work closely with the Evaluation Office at UNICEF and take a collaborative stakeholder engagement approach, convening dialogue and interaction with UNICEF’s Evaluation Office, the Evaluation Reference Group (ERG) and the Youth Advisory Group (YAG), as outlined above. This will ensure meaningful participation and engagement of key stakeholder groups.
- **Focus on feminist research principles**, intersectionality and gender equality and social inclusion (GESI): We put GESI at the heart of everything we do and operate on feminist principles. We will draw on our signature GESI expertise and bespoke tools, and voice and empowerment approach to assess progress around gender transformative programming and GESI mainstreaming within the selected programmes. We will ensure GESI is appropriately considered at each stage of the evaluation, that data collection reaches a diverse range of stakeholders and rights-holders, and that data is disaggregated by at minimum gender and age, and where safe, disability status and other identity

¹¹ UNICEF (2023). [UNICEF Evaluation of Impact Strategy and Action Framework 2022-2025](#).

characteristics. Feminist research principles have guided the constitution of the core research team, comprised of a group of eight women originated from the Global North and Global South and representing five nationalities as well as national consultants from each of the countries where case studies will be conducted (7 in total). This will ensure that a plurality of knowledge is drawn from. Knowledge embedded in beneficiaries will also be harnessed through workshops conducted with adolescent girls and boys. The team will pay attention to existing power dynamics and seek to mitigate them, whether they are based on national identity, knowledge or position (between interviewee and respondents for example), and will strive to conduct the evaluation in a non-extractive way. Finally, we will ensure that findings are shared back in user-friendly formats with participants by the CO and the YAG where appropriate.

4.5 Ethical approach and safeguarding

In addition, the evaluation will be guided by a set of ethical principles. These principles will be applied to the evaluation to minimize the risk of doing harm, while seeking to maximize the benefits of the evaluation. The principles will translate into practical measures to ensure confidentiality, informed consent, data protection, reduction of direct and indirect risks to interviewees, and safe and meaningful participation. This ensures that core principles such as Do No Harm and Leave No One Behind are adhered to and assure human rights, gender equality and equity considerations are built into the evaluation approach.¹²

All our research is informed by ethical and safety considerations for research and informed by the UNEG Ethical Guidelines, WHO’s ethical research guidance¹³, UNICEF Procedure on Ethical Standards in Research and Evaluation and UNICEF’s Ethical Research Involving Children, UNICEF Guidance on Gender Integration in Evaluation, as well as SDDirect’s Ethical Policy and our Child Protection and Vulnerable Adults (Safeguarding) Policy, which sets out our values and principles and describes how we meet our commitment to create a positive and safe environment for children and vulnerable adults who may be connected to our work. All our network consultants are required to read and confirm their commitments to these policies. All staff have received an induction on these policies.

The national consultants involved in data collection activities (see [Section 9](#) Team members roles and responsibilities) will attend an online workshop, which will cover all ethical principles and procedures in the evaluation. The training will include SDDirect’s safeguarding policies and code of conduct, including how to report suspected safeguarding concerns and how to act if a participant discloses situations of violence and/or abuse.¹⁴

5. EVALUATION FRAMEWORK

5.1 Evaluation questions

We present below the main evaluation questions. The full matrix outlining assumptions, indicators and main data collection methods and analysis is included in Annex 4: Evaluation Matrix.

EQ1: To what extent are the current Gender Policy and the GAP 3 conceptual framework well-designed, relevant, and coherent to respond to the various needs of all children, adolescents and women, national government priorities, UNICEF normative framework (i.e., CRC, CEWAD, CRPD, CCC) and the organization’s commitments to the UN-SWAP on gender equality? (coherence and relevance)

Sub-questions:

¹² See more information on how we will uphold our Ethical principles in Annex 2

¹³ See UNEG Code of Conduct for UN evaluation (2008) and WHO (2016) Ethical and Safety Recommendations for Intervention Research on Violence Against Women and UN Protocol on SEA

¹⁴ Please refer to Annex 2 for further detail on the ethical principles that will be applied through this evaluation.

- 1.1. To what extent are the current Gender Policy and GAP 3 informed by evidence and relevant to the various gender realities and needs of all children, adolescents and women across regions, particularly concerning gender and social norms? (relevance)
- 1.2. How well does the Gender Policy and GAP align to UNICEF's normative framework and the organization's commitments to UN-SWAP, and respond to partner government priorities to advance gender equality and equitably promote the rights of all children, adolescents and women in development and humanitarian settings? (coherence)

EQ2: How well are the gender-related organizational changes and performance enablers of the GAPS implemented across UNICEF policies, practices, systems, organizational commitments and accountability mechanisms at all levels? (coherence, effectiveness, efficiency and sustainability)

Sub-questions:

2.1 Enablers for gender responsive work

- To what extent have **gender equality programmatic approaches** for transformative results been systematically **integrated into the Strategic Plan and regional strategies as well as in UNICEF programming cycles at the country levels**, including country programme documents, emergency preparedness, response and recovery plans? (effectiveness)
- To what extent has UNICEF harnessed data, research and evaluation to inform evidence-based policies and programmes and promote gender equality and adolescent girls' empowerment? (effectiveness, efficiency)
- How successfully has UNICEF mobilized resources and met the UN-SWAP standards on gender equality and the empowerment of women resource allocation benchmark of 15 per cent for gender-related work? (effectiveness)
- To what extent are human resources capacity and systems adequate to support the implementation of the GAP?

2.2 Partnerships

- To what extent have strategic partnerships been identified and leveraged to advance gender equality and the empowerment of girls and women? Have partnerships with girl-, youth- and women-led organizations and networks been prioritized? (effectiveness, efficiency)
- How effective has UNICEF been in supporting the inclusion of a gender perspective within national systems and structures, and the absorption of UNICEF's gender-related initiatives by partners to promote sustainability? (sustainability)

2.3 Accountability and leadership

- How well have gender-responsive monitoring systems and accountability mechanisms been utilized to improve learning and accountability at all levels? (effectiveness)
- How precise are the accountability mechanisms to ensure that gender programmatic results are integrated at all levels?
- To what extent does senior leadership demonstrate their commitment to gender Equality? (effectiveness)

2.4 Staffing and culture

- To what extent is gender equality considered "the responsibility of everyone at UNICEF – at all levels and in all offices of the organization", as GAP 3 suggests,

and do staff and management feel ownership, responsibility for, and engage in its implementation?

- How successfully has UNICEF as a workplace managed to promote gender equality when it comes to staffing and the inclusion of diverse staff in all sectors, including senior management; recruitment, hiring and promotion processes; and zero tolerance for sexual misconduct? (effectiveness)

EQ3: To what extent have UNICEF’s programmatic results for gender equality throughout the life course been met across all five Goal Areas, particularly to advance adolescent girls’ leadership and well-being, in both development and humanitarian settings? (effectiveness, impact)

Sub-questions:

- 3.1 In looking at examples where UNICEF has significantly contributed to gender equality across all **five Goal Areas**, **what results** have been achieved?
- 3.2 What results have been achieved in advancing **adolescent girls’** leadership and well-being? (impact)
- 3.3 To what extent has UNICEF contributed to producing **transformative changes in gender norms and systems**, and to tackling structural power dynamics, whether intended or not? (impact)
- 3.4 What have been the enabling (and hindering) factors and processes, both within and outside UNICEF? (effectiveness)
- 3.5 What **lessons** can UNICEF learn for broader application in programme design and implementation? (effectiveness)

The table below illustrates our suggested amendments to the original evaluation questions proposed in the ToRs and justifications for the change

Table 4: Proposed amendments to the evaluation questions in the ToRs

Evaluation question defined in the ToRs	Proposed amendment	Justification
Question 1		
Evaluation Question 1	We propose limiting the analysis of relevance and coherence of the GAPs to the GAP 3 under this question.	The relevance and coherence of the GAP 2 was analysed in the previous evaluation in 2019, and we will use the findings and recommendation to check if they were implemented.
Sub-question 1.c	This sub-question question will be explored under sub-question 1.1 which becomes: <i>To what extent are the current Gender Policy and GAP 3 informed by evidence and relevant to the various gender realities and needs of all children, adolescents and women across regions, particularly concerning gender and social norms? (relevance)</i>	To rationalize and focus the evaluation questions, we grouped the topics exploring relevance in one single question.

Evaluation question defined in the ToRs	Proposed amendment	Justification
Sub-question 1.e	<p>This question will be explored under SQ2.1:</p> <p><i>To what extent have gender equality programmatic approaches for transformative results been systematically integrated into the Strategic Plan and regional strategies and in UNICEF programming cycles at the country levels, including country programme documents, emergency preparedness, response and recovery plans? (effectiveness, coherence)</i></p>	<p>There was an overlap between both questions so for conceptual clarity and to avoid repetition, we propose to combine them.</p>
Question 2		
Evaluation Question 2	<p>We grouped the questions under EQ 2 focused on gender-related organizational changes under four categories:</p> <ul style="list-style-type: none"> • Enablers for gender responsive work • Partnerships • Accountability and leadership • Staffing and culture. 	<p>These categories mirror the “inputs” depicted in the ToC and will be analysed together.</p>
Sub-question 2.g	<p>We have rephrased part of this sub-question and pulled it out as a separate sub-question under sub-section 2.3 <i>Accountability and Leadership: To what extent does senior leadership demonstrate their commitment to gender equality? (effectiveness)</i></p>	<p>The new question allows for a more nuanced analysis.</p>
Question 3		
Evaluations questions 3, sub-question 3.a and sub-question 3.b	<p>The evaluation will adopt UNICEF definition of impact¹⁵, but it is beyond the scope of this evaluation to measure higher-level impacts.</p>	<p>The evaluation budget and scope programmes are not sufficient to ascertain the impact of UNICEF programmes on adolescent girls through primary data collection at the global level, so we will rely on the analysis of secondary data from the GAP indicators. In addition, it is not possible to measure the impact of GAP 3 after just two years of implementation.</p>
Question 3.c	<p>We propose to amend the question to: To what extent has UNICEF contributed to producing transformative changes in gender norms and systems, and to tackling structural power dynamics,</p>	<p>The request to broaden the scope to structural root causes was made by the ERG.</p>

¹⁵ As described in Section 4.3

Evaluation question defined in the ToRs	Proposed amendment	Justification
	whether intended or not? (impact)	
Sub-question 3.d	Partnerships will be tackled through SQ2.2: <i>To what extent have strategic partnerships been identified and leveraged to advance gender equality and the empowerment of girls and women? (effectiveness, efficiency)</i>	Measuring the relative contribution of UNICEF vis-à-vis partners in implementing interventions and identifying impact results would warrant a study in itself, and the current evaluation cannot tackle this topic within the time and resources imparted. We will however look at the effectiveness of those partnerships, with a focus on partnerships with girl-led, youth and women-led organizations.
Sub-question 3.e	We propose to amend the question to: What lessons can UNICEF learn for broader application in programme design and implementation? (effectiveness)	The request for clarification was made by the ERG.

5.2 Evaluation Theory of Change

The theory of change (ToC) was designed during the inception phase of the evaluation through a consultative and participatory approach (see details on the approach in [Section 4](#)). The team discussed how to define the problems, inputs, outputs and outcomes for the purpose of the evaluation but also importantly, time was spent identifying the assumptions along the different pathways of change.

These assumptions will be tested during data collection and have informed the development of the data collection tools. They are also included in the evaluation matrix, organized per research question, and complemented with the data sources and analysis that will be conducted to test each of them (see Annex 4: Evaluation Matrix). The figure below presents the ToC as well as the assumptions. An accompanying narrative presenting the logic of the various change pathways is presented in Annex 3.

Figure 2: Evaluation theory of change

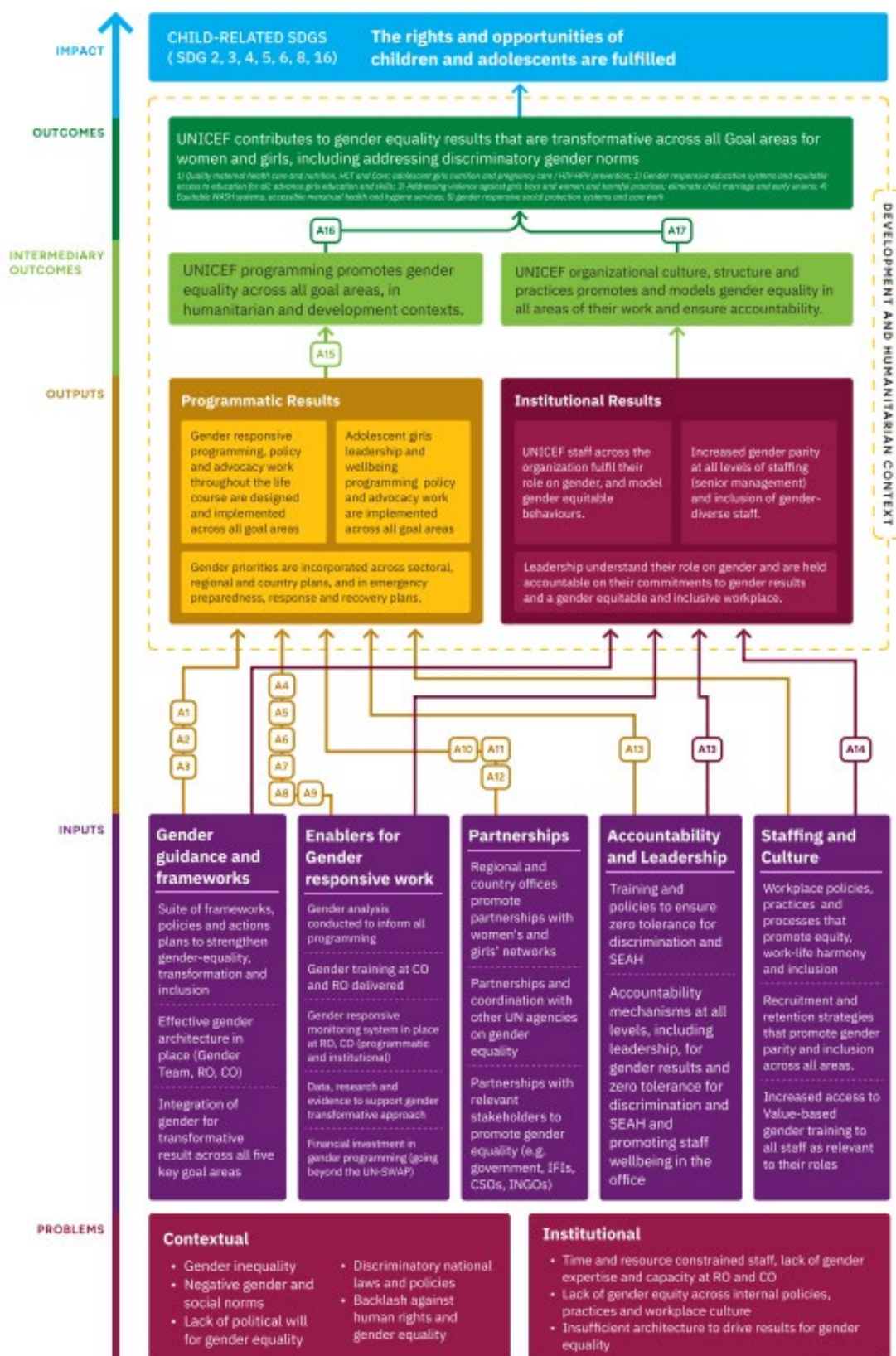


Figure 3: Evaluation theory of change assumptions

Assumptions



5.3 Evaluation matrix

The full version of the evaluation matrix presents the evaluation questions, a short narrative of our understanding of each question and key elements to pay attention to during the evaluation, based on preliminary consultations and information collected during the inception phase and recommendations from the previous evaluation. Under each question, we have mapped out the assumptions we collectively defined when we developed the ToC, as well as the indicators and sources of data and types of analysis we will conduct.

6. METHODOLOGY

6.1 Summary of overall methodology and documents produced

For this evaluation, we will apply a participatory and mixed-methods approach drawing on primary and secondary data to answer the evaluation questions.

We will sequence the analysis which will involve exploring global data trends first, to identify emerging issues to be explored in greater depth during the primary data collection in countries in which UNICEF operates and at the global level, and across staff levels as well as outside the organization during a second phase.

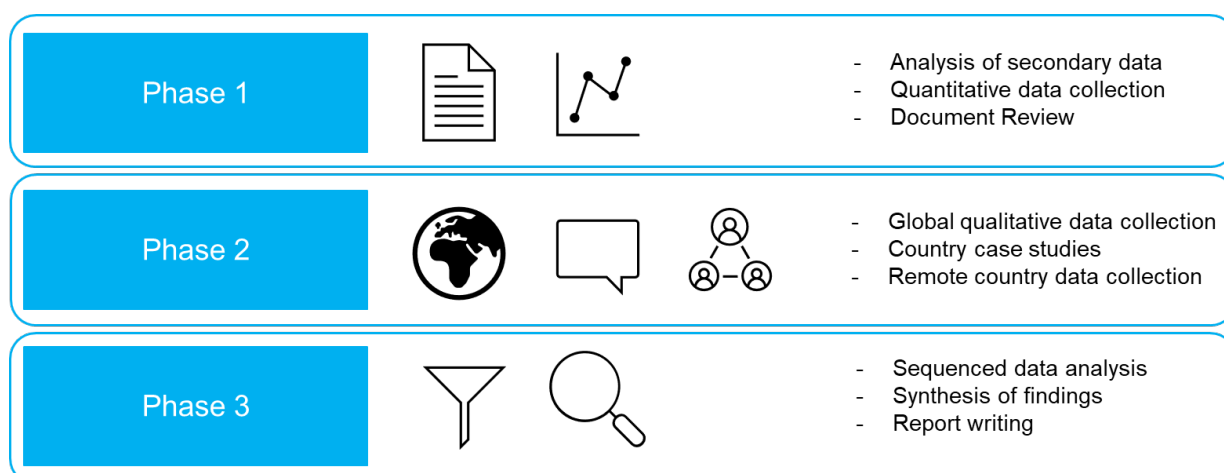
The evaluation methodology will be done in three phases, as outlined below, and will result in five deliverables (inception report, interim report¹⁶, pilot and full case study reports, final evaluation report and dissemination products – as described in section 7 below).

1. **Phase one: Analysis of global data and evidence.** The initial phase of the evaluation will take a global focus to assess the implementation of the institutional enablers and programmatic results as well as emerging trends through secondary data analysis.
2. **Phase two: Primary data collection.** This phase will consist of collecting primary data, which will be informed by emerging results from phase 1. Three sets of primary data will be collected concurrently, and the data collection period will overlap with phase 1 to ensure efficient delivery of the evaluation¹⁷:
 - 2.1. A staff survey KIIs with HQ and regional staff to explore organizational and operational themes (see details in section 2.2).
 - 2.2. A comparative study involving three different organizations will be conducted. It will provide examples of good practice and practical ways to respond to some of the challenges experienced in UNICEF of integrating gender, and feed into the discussions around final recommendations (see details in section 2.2).
 - 2.3. At country level, primary data collection will build on the Interim report findings to explore the application of the GAP 3 and its contribution to workplace culture and programmatic results within a sample of country offices. Short case studies reports will be produced for each case study conducted (we are proposing a maximum of seven case studies (six complete and one top-up) across the priority and back-up countries selected during inception (see details in section 6.3).
3. **Phase three: Synthesis and report writing.** This phase will focus on analysis and triangulation both across and within categories of data sources (global data, case countries findings, survey, comparative analysis) to contribute to the overall findings, conclusions and recommendations. A final evaluation report will be produced (60 pages)

¹⁶ This will include a descriptive analysis of the survey results and of the GAP standard index indicators

¹⁷ We will seek to finalize the global data and evidence before commencing primary data collection in countries and through the survey.

Figure 4: Evaluation methodology phases



6.2 Analysis of global data and evidence

6.2.1 Desk review (documents and data)

A comprehensive desk review will take place, to complement the initial document review that was undertaken as part of the inception period.

Document review

The evaluation team has already reviewed a range of UNICEF documents during the inception phase,¹⁸ we will extend this review to a wider range of documents related to the institutional change pathway and the programmatic change pathway articulated in the ToC. This would include, but not limited to:

- **Institutional change pathway:** Human resources policies and frameworks, annual reports, board reports, strategic plan performance review, accountability framework, policy documents and papers relating to institutional enablers (for example, the Background Paper Series).
- **Programmatic change pathway:** Gender-focused evaluation reports, relevant research produced on gender equality, GAP data companion reports, strategic plan annual reports, a sample of Gender Programme Reviews and related Country Programme Documents (for a max of 20 countries), as well as humanitarian gender analysis and related programming documents and plans.

The documents will be coded thematically against a matrix that aligns with the evaluation question assumptions. During this process, the team will also identify key outcomes within the documentation that align with the evaluation ToC.

A gender analysis of the key guiding frameworks will be conducted to assess alignment with the Gender Policy vision. The analysis of the documents related to the programmatic change pathway will be done against the GESI continuum.

To aid this process, UNICEF Evaluation Office has proposed the support of a data scientist, who will utilize Natural Language Processing (NLP) to support the synthesis of relevant documents and large volumes of qualitative documents, including those describing results and gender analysis, using machine learning. Our expectation is that this will enable us to assess the scale at which various gender results, categorized against the GESI continuum, are reported.

More specific data sources are presented in the evaluation matrix.

¹⁸ The initial desk review informed the evaluation matrix, the approach and methodology

Data Review

UNICEF has a range of data sources that will be analysed to draw out global trends across:

- The implementation of the institutional enablers of the GAP
- The workplace culture and progress on organizational culture change
- The gender equality results that are transformative across the five Goal Areas.

To do this, we will draw on the GAP Standard index, the global annual staff survey, the pulse check on workplace culture, the indicators reported in the GAP data companion and the gender spend data including the gender marker (GEM) and the activity tag.

Initial analysis of quantitative data will focus on drawing out key trends across the period 2019-2023 where possible according to a cross over of the indicator sets, as relevant to the institutional and programmatic change pathways.

Institutional change and enablers data

The evaluation team will draw on various data sources to understand the trends in relation to the institutional enablers and the change pathway. A key data source here will be the GAP standard index which provides an overall score on GAP institutional enablers for country offices.

We have identified indicators from the GAP standard index that represent the “how” and the “who” institutional enablers dimensions, and mapped their consistency over GAP 2 and GAP 3. To see the full list of proposed indicators for this analysis in Annex 12: Selected indicators for quantitative data analysis.¹⁹

To supplement this, the evaluation team will also run analysis on the Global Staff Survey and Pulse Check surveys data collected by UNICEF to understand perceptions of staff in relation to workplace culture and accountability mechanisms for gender equality and work-life harmony, disaggregated by gender.

Programmatic results data

There will be two main data sources considered: indicators reported in the GAP data companion and gender-related budgetary data.

The indicators reflected in the GAP data companion and section 9 of the GAP standard will form the main data sources for programmatic results data analysis. Trend analysis will be completed for the indicators in the GAP data companion that have been consistently reported between 2019-2023 (across GAP 2 and GAP 3). This will be completed by a sub-set of other GAP 3 programmatic indicators reported as part of the Data Companion in 2022 and 2023. The purpose of this is to understand the trends related to gender equality results that are transformative across the five key goal areas, through the life-course and for adolescent girls. At global, regional and country level the Evaluation Team will also conduct budgetary analysis of gender-related budget allocations and spending patterns across goal areas, life-course and adolescent-girl focused programmes and interventions, and humanitarian spending. Areas that will be covered include:

- Evolution of gender expenditure since 2019 – looking at transformative and integrated expenditures
- Evolution of gender transformative expenditures in humanitarian settings
- Evolution of funding sources towards gender equality activities.

Further analysis

A regression analysis will be conducted for 2022 and 2023 using the GAP standard index and key

¹⁹ The list will be finalized through further conversations with UNICEF.

programmatic results indicators (from section 9 of the GAP Standard) with the purpose of exploring the relationships between the implementation of the institutional enablers and the quality of gender equality results for women and girls.

6.2.2 Primary data collection

Primary data will be collected during the case studies and via an all-staff survey. We will use interviews and/or focus group discussions with staff from HQ, ROs and COs and evidence from the comparative analysis exercise.

Survey

We propose to conduct an all-staff survey disseminated by the Evaluation Office to HQ, ROs and CO's. The purpose of the survey is to collect data on the views of staff on the implementation of GAP institutional enablers, both on organizational change enablers to deliver programmatic results (including the integration of gender equality in programme analysis and planning, use of monitoring data of gender equality, human and financial resources availability, partnerships) and on performance enablers focusing on gender in the workplace. The survey will also investigate the contribution of these enablers to improving UNICEF's organizational culture and its programmatic results on gender equality. The survey will also enquire about the challenges respondents encounter in their day-to-day work on gender equality and their suggestions for improvements. The survey questionnaire will be kept short and questions will be tailored based on the role of the respondents using skip logic (e.g., gender staff vs senior management) and will not duplicate questions collected through the Global Staff Survey/Pulse Check. The survey will be kept open for 3-4 weeks and will inform the qualitative data collected through the subsequent interviews and focus group discussions during the country visits. Results will be triangulated with the other qualitative data.

Descriptive analysis of the survey response data will be conducted in Excel and we will apply sub-group analysis to ensure the application of an intersectional lens, whereas qualitative analysis will apply a thematic analysis method to draw out key trends.

Key informant interviews and focus group discussions

A series of key informant interviews (KIIs) and focus group discussions (FGDs) will take place with UNICEF staff at the three levels of the organization and partners at country level. The sampling will take a purposive approach drawing from a key stakeholder list provided by the UNICEF Evaluation Office, this will be complemented by a snowballing approach. KIIs will follow a topic guide covering the relevant sub-questions and will be informed by the findings of the data review and survey. These interviews will be semi-structured.

Table 5: Provisional list of KIIs/FGDs (to be finalized with the Evaluation Office)

Respondents	Number of KII/FGD	Key themes
UNICEF Division of Human Resources and Culture and Diversity team	1	Workplace policies on gender and SEAH. Gender parity and measures to prevent gender bias and enhance women's wellbeing at work. Accountability of managers on gender equality. Organizational culture.
UNICEF Finance Team	1	Gender expenditure and GEM. Funding sources for gender work.
UNICEF DAPM	1	Knowledge management. Alignment of GAP with other UNICEF frameworks (internal coherence). Support with monitoring and data analytics.
UNICEF Programme Group	1	To explore how gender was integrated across the sectors, progress made and remaining challenge (1FGD).

UNICEF ADAP	1	Inclusion of adolescent girls in the GAP, its effectiveness, relevance and impact.
UNICEF Gender team	1	The implementation of the GAP, the impact on programme results, what has worked, challenges, hampering and enabling factors, levels of influence (1 FGD with the whole team).
UNICEF EMOPS	1	Relevance of the GAP to humanitarian context; programming results; successes, challenges and hampering factors specific to humanitarian contexts; integrating gender in the peace/development nexus; institutional enablers specific to the context.
UNICEF Leadership (ED, DEDES)	3	Strategic vision for gender equality within the organization and in its programmes; perception of the GAP results and challenges (3 KIIs of 30 minutes).
UNICEF Innocenti/Global Office of Foresight and Research	1	Use of research and data to inform programmes. Gender transformative agenda.
UNICEF Regional Management	7	Implementation of GAP programmatic priorities Gender equality results; regional contextual hampering and enabling factors for integrating gender equality (7 KIIs of 30 minutes with Deputy Regional Directors).
Members of the Executive Board and UNICEF Partners (including INGOs, UN agencies)	4	Executive Board members: Strategic vision for gender equality within the organization and in its programmes; perception of the GAP results and challenges. UNICEF partners: Perception of effectiveness of partnership with UNICEF for promoting gender equity for transformative results Staff capacity on gender equality.

All KIIs will be analysed using a thematic analysis approach²⁰ to identify key trends and patterns. Both inductive and deductive coding will be used against a base coding framework aligned to the evaluation matrix, including new thematic codes as trends are identified during analysis.

6.3 Comparative analysis study

A comparative analysis study will be conducted with up to three UN, multi-lateral and bi-lateral/donor agencies. The purpose of this study is to compare differing approaches to the integration of gender across similar organizations with the aim of drawing out lessons and good practice examples to inform the evaluation recommendations.

The study will draw heavily on published data by selected agencies, as well as other key documents such as Multilateral Organizational Performance Assessments.

To select the sample of organizations, the evaluation team developed a selection-criteria, informed by qualitative data from the KIIs. The proposed selection criteria include:

- Similar target population (including adolescent girls and young women, adolescent boys and young men);
- Organization with recognized track record on gender transformative approaches and innovation on promoting gender equality and equity;

²⁰ Thematic analysis involves identifying and analysing patterns or themes within qualitative data. It is a flexible and adaptable approach that can be used to analyse data from various sources, such as interviews, focus groups, and open-ended survey responses.

- Organization with similar programme size as UNICEF; and
- Humanitarian operations/protection work.

A main challenge anticipated in this study is securing participation of agencies to another organization's evaluation. To maximize the chances of obtaining our desired sample of agencies, we will adopt an iterative approach to identifying respondents by first contacting three organizations within a long list of suitable candidates, and contacting additional ones depending on interest in participating. In order to encourage participation, we will clearly communicate to potential participants the benefits of the exercise (e.g. cross learning and showcasing good practices), as well as on measures to ensure a safe space for discussion. These include maintaining confidentiality and anonymity of KII and outcomes of group discussions; presenting results at an aggregate level; drawing learning without identifying specific agencies; ensuring that if a specific agency is quoted it will only be based on already published material and; validating the draft report with participating agencies.

Qualitative data collected during the inception phase identified a number of organizations for consideration for inclusion in the comparative study. The table below illustrates how those organizations meet the proposed criteria.

Table 6: Proposed organizations for comparative analysis in relation to criteria

Criteria	UNFPA	UNDP	UNHCR	WFP	UN WOMEN	Plan UK	Gates Foundation	Save the Children UK	IRC	OXFAM	Global Affairs Canada
Similar target population											
Organization with recognized track record on gender											
Organization with similar programme size											
Humanitarian operations/protection											

Based on this, organizations that meet at least three of the four proposed criteria are: UNFPA, UNHCR, WFP, Plan International UK, Save the Children UK, OXFAM and Global Affairs Canada. From those seven suitable organizations that form our long list, it is proposed to first narrow down to three, bearing in mind that if some of those were to decline to participate in the exercise, another organization on the long list would be contacted. For the first set of three, we propose to include two UN organizations: UNFPA and UNHCR and one donor: Global Affairs Canada (GAC). While UN organizations offer the most comparable institutional context to UNICEF, including an organization implementing bold, innovative gender transformative approaches at scale (GAC) allows diversifying the sample. While GAC's nature and operational set-up differ significantly from UNICEF, respondents have highlighted that this organization meets the criteria we have identified and may offer key learning opportunities on institutional and programmatic enablers for gender equality. More precisely, we anticipate interesting learning from the way feminist principles are

enacted in the institutional culture and in the workplace, in the internal operations (including Human Resources and accountability mechanisms), as well as promoting feminist funding to women’s and girls’ organizations.

The comparative analysis process will involve the following phases:

- Identifying organizations and confirming their interest in participating in the comparative analysis on gender transformative approaches;
- Collating and analysing relevant secondary data;
- Conducting individual or group interviews with each of the 3 organizations to understand their processes and identify additional documentation;
- Inviting the participating organization to a session to discuss key findings from this exercise to refine findings; and
- Sharing the comparative analysis report with them for validation.

We will interview one to two respondents per agency, at HQ level, covering responsibilities for gender integration in programmes and in the workplace. A semi-structured questionnaire guide will be designed, covering the thematic areas and topics that are presented in Annex 5: Data collection topic guide – Comparative analysis.

To supplement the information shared during the interviews, we will ask agencies to share the following documentation to substantiate findings: gender policies and strategy documents, theory of change, monitoring and evaluation frameworks, relevant reports including HR reports, evaluations and assessments relating to programme and institutional integration of gender.

Table 7: Proposed KIIs/FGDs for comparative analysis

Respondents	Number of KII/FGD	Key themes
Institutions participating in the comparative analysis	6	Policy and programme framework on gender Gender architecture and HR Accountability mechanisms on gender in the workplace Level of resources for gender and how it is calculated Contribution of institutional enablers to gender results Gender and other cross-cutting priorities
FGD with the same	1	Validation of emerging findings and discussion of lessons learned

The document review and KIIs conducted as part of the comparative study will be coded against a tailored coding framework, allowing for comparisons to be drawn across the frameworks and approaches used by the organizations to integrate gender, their implementation and their contribution to programmatic and institutional results. Secondary and primary data relating to each participating organization will be triangulated under each of the dimensions of analysis (see Annex 5 for more detail): gender transformative approach; gender in policies and programming processes: gender analysis, M&E and use of data; human and financial resources; partnerships; accountability and leadership; gender equality in the workplace; organizational culture; gender and other cross-cutting priorities; key learnings).

Participating organizations will be invited to a 90-minute online workshop where emerging findings will be fed back in the form of a slide deck to generate discussions and learning. A short standalone report on the comparative analysis will be developed and shared with participating agencies for

their feedback.²¹ The report will consist in a simple visual summary of the key results, as well as an analytical section describing learning and good practices from the agencies under each dimension.

6.3 Case Study Methodology

The evaluation will include a maximum of seven country level case studies.²² The first case study will be a pilot conducted by the Team Lead, during which the data collection tools will be tested, and subsequently revised to guide data collection in the remaining countries. An updated protocol will be developed drawing learnings from the pilot and relay any changes to the implementation of the methodology.

The case study methodology will apply aspects of appreciative inquiry²³ to understand organizational change in line with the gender strategy, the GAPs and the results achieved for adolescent girls and young women and through integrating gender across sectors.

To cover a representative cross-section of UNICEF's work, the Evaluation Team proposes undertaking one country case study in each region where UNICEF works, this will include six in-depth country case studies and one 'top-up case study' building on the recently completed Formative Evaluation of Gender Transformative Programming through Investment in Adolescent Girls' Leadership in Eastern and Southern Africa (2023), conducted by SDDirect. The selection of countries has been made during the Inception Phase identifying COs that have effectively brought about institutional shifts in gender work, or countries that have reported gender equality results that are transformative according to the UNICEF reporting mechanisms. This will help to capture learning of what works across different contexts, in line with our appreciative inquiry approach.

To avoid positive bias, the team will supplement these case studies with data collection (survey and additional KIIs) from five country offices which do not fulfil the appreciative inquiry criteria. This will help capture challenges and hinderances to implementing the Gender Policy and GAP across UNICEF contexts of work.

6.3.1 Case study selection

A systematic and purposive approach to country selection was taken during the inception phase, drawing on secondary quantitative data and primary qualitative data (a short survey was shared with Regional Gender Advisers and Regional Evaluation Advisers).

To identify high performing countries in relation to gender integration, the team considered the following: a GAP standard score of above 70%, reports of transformative gender equality results and the existence of a Gender Flagship Program. This was complemented by qualitative data collected from Regional Gender Advisers and Regional Evaluation Advisers, sharing examples of good practice from across the country offices.

To ensure that this sample was representative of a range of UNICEF specific and broader contextual factors,²⁴ a criteria-based approach was taken. Whereby we sought to identify a sample that was representative of:

- UNICEF's regional coverage;
- UNICEF's humanitarian and development portfolios;

²¹ Findings presented in the comparative study report will be high-level and compiled across the organizations without identifying them individually in order to respect confidentiality (5-10 pages). Where relevant, specific examples citing an agency may be presented from published documents. This format will be explained at the onset of the process so that respondents are clear about how their input will be presented in the evaluation report.

²² As previously noted our ability to deliver seven case studies is dependent on COs' willingness and capacity to be involved during the timeframe of this evaluation.

²³ See details on this in section 4.1: analytical approach.

²⁴ Please note that to avoid evaluation fatigue, a range of countries that have been heavily evaluated in the last three years were removed from the sample.

- UNICEF’s work across the five key goal areas;
- Country Office budget, size and gender spend;
- Low, middle and high-income country contexts;
- Humanitarian/FCAS settings; and
- Broad gender equality context in the country.

A two-stage process was applied using these criteria to identify seven country offices for case study implementation.²⁵ The final list is shared below, with a rationale, and the full process for selection is outlined in Annex 11: Country selection matrix.

Challenges in Methodology and Caveat:

In six of the seven regions, our selected country offices have agreed to participate in the evaluation as a case study. In EAPRO, we have struggled to identify a country office that aligns to the methodological approach and whom is willing/able to participate.

For the EAPRO region, we approached four country offices that met the selection criteria and aligned to the methodological approach however none were able to participate in the evaluation. The Multi-Country Office, Fiji, has been suggested by the UNICEF Gender Team as an option. The Fiji office does not meet the selection criteria selected and sits outside of the appreciative inquiry approach. We are in discussion with the Fiji Office about their participation, and determining the focus of the case study in the coming weeks.

In the case that this case study does not go ahead, we propose to reallocate the LoE to two other data collection activities under the evaluation:

1. Data collection on challenges in country offices not meeting GAP standard target: Increase remote data collection with country offices that did not meet the GAP standard target, we could include Fiji within this sample. This could be through the increase in number of country offices or the number of interviews with those country offices.
2. We would increase global data collection to reach a wider range of stakeholders.

²⁵ Six full in-country case study countries and one top-up case study country

Table 8: Proposed case study countries

Region	Final Country Selection	Rationale	Alternative Country Option(s)
ECA	Bosnia and Herzegovina	Development context, presence of Gender Flagship Program, reports gender equality results that are transformative, scores >70% on GAP Standard Score. UMIC with low GII score. Small office size with smallest regional CO budget and gender spend.	Tajikistan
ESARO	Mozambique ²⁶	FCAS context, presence of Gender Flagship Program, reports gender equality results that are transformative, doesn't score >70% on GAP Standard Score. LIC with mid-range GII score, large office size and CO budget, highest gender spend from regional shortlist.	No-back up
LACRO	Peru	Development context, presence of Gender Flagship Program, reports gender equality results that are transformative, doesn't score >70% on GAP Standard Score. UMIC with low GII score. Medium sized CO, mid-range CO budget but highest gender spend from regional shortlist.	Dominican Republic
MENA	Egypt	Development context, presence of Gender Flagship Program, does not report gender equality results that are transformative, doesn't score >70% on GAP Standard Score. LMIC with mid-range GII score. Large sized CO, lower CO budget and gender spend from regional shortlist.	Jordan
ROSA	Bangladesh	Development context, presence of Gender Flagship Program, reports gender equality results that are transformative, doesn't score >70% on GAP Standard Score. LMIC with mid-range GII score, large office size with large CO budget and high gender spend.	Nepal Bhutan
WCARO	Democratic Republic of Congo	FCAS context, no Gender Flagship Program, reports gender equality results that are transformative, does not score >70% on GAP Standard Score. LIC with highest range GII score from shortlist, large office size with highest CO budget and gender spend from shortlist.	Guinea

²⁶ Mozambique featured as a case study country in SDDirect's Formative Evaluation of Gender Transformative Programming through Investment in Adolescent Girls' Leadership in Eastern and Southern Africa (2023), and can therefore be the top-up case study for this evaluation. See Section 6.3.5. There is no available top-up for this country that meets the required criteria for a top-up study.

6.3.2 Objectives of the case studies

The case studies will explore the implementation of the GAPs and Gender Policy at the country level, the contribution to shifts in workplace culture and accountability and interrogate their contribution to improved programmatic results for women and girls.

6.3.3 Case Study Preparation

In each case study country, a core Evaluation Team member and a National Consultant will work together to collect data and complete the evaluation. The case study process will begin with a preparatory phase during which the following actions will take place:

- Recruitment of National Consultants with gender and evaluation expertise.
- All travel planning and due diligence is completed.
- All data collection tools will be developed and/or adapted according to purpose, context, language and accessibility needs.
- All participating case study countries are confirmed with the support of UNICEF EO, with interviews and workshop dates scheduled accordingly by the National Consultant.
- All National Consultants and core team members complete half-day orientation training to review the evaluation questions, approach and methodology and safeguarding policy and code of conduct. This will be conducted up to three times to ensure all National Consultants can participate across different time zones.
- Any final adaptations to the case study methodology are made based on findings from the analysis of global data and evidence, and meaningful engagement of the YAG. A two-hour workshop with the YAG will take place providing an opportunity for the group to engage with the case study design and influence implementation.

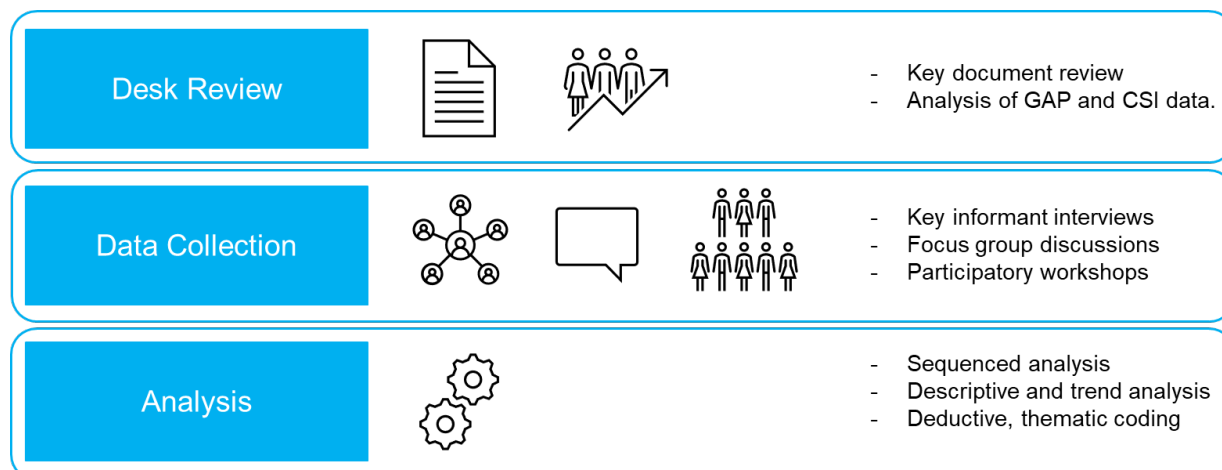
Each Evaluation Team member and partnered National Consultant will hold the relationship with the country office. Before data collection begins a kick-off meeting will be held with key CO staff to introduce the evaluation, the case study methodology, timeline and deliverables. Space will be created for the country office team to provide feedback that will aid contextualization of the case study methodology. Data collection instruments will be refined, translated into local languages and pre-tested, and revisions made as required.

6.3.4 Case study methodology

The case study will take a mixed methods approach, drawing on both primary and secondary data sources. These case studies will draw heavily on participatory and qualitative methods, such as:

- Desk review: An in-depth review of key programme documentation and quantitative data, including mid-term assessments, evaluations, ToCs, process reports, financial data and programme results data, and deeper analysis of country specific global survey data.
- Key informant interviews with UNICEF staff and stakeholders (e.g. women's networks, other UN agencies, government and other civil partners).
- Focus group discussions with adolescent girls and boys involved with UNICEF utilizing participatory methods.
- Participatory workshop with UNICEF CO staff to explore attitudes, perceptions, norms, barriers and enablers internally and externally.

Figure 5: Full Case study methodology



Desk Review

For each case study, a desk review will be conducted, consisting of both a document review and a review of existing quantitative data.

- Document review: Documents will be sourced from UNICEF EO and the CO teams, as well as through interviews with stakeholders identifying further documentation for review. These documents will include: programme documents, annual reports, programme evaluations, gender analyses, etc. A consistent coding framework will be designed in line with the EQs and sub-questions, and thematic coding will be applied using Dedoose.
- Results data analysis: Existing UNICEF data available on the GAP Standard, the GAP data companion, gender spend data and programme results data will be analysed and trends between the different factors examined.

This desk review will take place once country selection is confirmed and ahead of primary data collection to further inform primary data collection tools and approaches. The findings from the desk review will be triangulated with the primary data through the case study.

Key informant interviews

Key informants will be identified through purposive sampling informed by a stakeholder mapping, programme documents and initial consultations with UNICEF COs, this will be complemented by snowball sampling. An average of 20 interviews per country is anticipated, with some countries requiring more and others less depending on the size of country programming. The interviews will take place both individually and in groups, using a semi- structured format.

We will develop a set of semi-structured interview guides which can be adapted based on the stakeholder type and their role within or relationship to UNICEF (see [Table 7](#)). The interview guides and the topics they will cover will be primarily informed by the EQs, evaluation ToC and the evaluation matrix. There will be a consistent set of high-level themes or questions that will be explored to ensure nothing is missed; however, scripts will be adapted to the needs of each stakeholder group and probes will be used to further explore relevant information. These scripts will also be adapted to the different country contexts and have space for questions to explore outcomes identified through the initial document review.

Table 9: Proposed stakeholders for in-country KIIs

In-country stakeholders	Purpose	Average no. of people/ country

UNICEF CO staff (includes programme staff, gender focal points, HR staff and management)	To gather information on understanding and implementation of the Gender Policy and GAP, as well as results, enablers and hindering factors and whether this has contributed to gender equality results that are transformative.	8-10 (KIIs)
Other UN agencies	To explore the scope and nature of inter-agency collaboration and the extent to which it contributes to better gender equality results, as well as their perceptions on UNICEF's contribution to gender equality and their comparative advantage.	3
Implementing partners (inc. Women's and girl's networks and other civil society partners)	To explore the nature of partnerships between UNICEF and civil society, examining the scope and quality of these partnerships and their contribution to gender equality results. Drawing out enabling and hindering factors for partnership.	6
Government stakeholders	To examine the political will to support gender equality and explore their perceptions of UNICEF's contribution to gender equality. Where they have worked in partnership with UNICEF, reflection on the nature and success of the partnership will be discussed.	2
Donors	To explore the role of donors in supporting gender equality and their perceptions of UNICEF's contribution to gender equality.	2

Focus group discussions

In each country, two focus group discussions (FGDs) will take place with adolescent girls and boys who have participated in UNICEF programmes. These FGDs will be facilitated by the National Consultant and Evaluation Team member to gather and explore stories of change in relation to the impact of UNICEF work.

The focus groups will be adapted to the participants and employ participatory and fun activities to build trust and rapport among the group and the facilitator. We will be exploring where and how adolescents and young people are able to exercise agency and decision making, in what aspects of their lives, how that has changed over time and how has UNICEF contributed to that, and what impact any increased agency has had on their lives. This will complement the appreciative inquiry approach to the case studies and seeks to highlight any changes or impact that UNICEF has contributed to, and uncover the pathways that led to them, as valuable lessons for future programming.

Implementing partners and/or UNICEF country and field office staff will support the logistics and planning for these discussions. Accessibility requirements will be identified in advance to ensure youth with disabilities can also meaningfully participate, and budget is available for reasonable accommodation and / or sign language interpreters as necessary.

The focus groups will be held in single-sex groups to ensure a safe space for reflection and to maximize participation among both girls and boys. A non-coercive, opt-in approach will be employed to select young people to participate in the FGDs.

The sampling approach will be purposive, with support from UNICEF Country Offices and relevant implementing partners to identify adolescents that have been impacted by UNICEF programmatic engagement or that engage with UNICEF through advocacy or advisory roles. Letters of consent and assent will be shared with young people and their parents/carers prior to any conversations,

and all young people (and their parents/carers) will have the right to withdraw (or withdraw their children) at any point during the process.

Table 10: Proposed stakeholders for in-country FGDs

In-country stakeholders	Purpose	Average no. of people/ country
Adolescents and rights holders	To explore the direct and perceived impact of UNICEF on youth. Drawing out enabling and hindering factors for gender equality work that is transformative.	10 – 16

Participatory workshops

We will convene two three-hour, single-sex workshops for 8-12 UNICEF staff participants per case study country (sex-disaggregated). A purposive sampling approach will be taken to identify UNICEF staff from across departments and grades in the Country Office. The purpose of this workshop is to explore the institutional change pathway, unpacking awareness, attitudes and perceptions towards gender equality within their programmatic work and within the organizational culture at UNICEF.

An adapted “looking in, looking out” or similar approach and methodology will be used for this workshop.²⁷ Through the application of this methodology, we will work with the group to explore values and attitudes towards gender equality (in programmatic work and the workplace), the culture in the office, how gender impacts their experience at work and examine power dynamics. These are complex subjects to explore. To ensure the safety and trust of participants, we will implement same-sex workshops, employ group work, begin the sessions with trust-building exercises and have regular check-ins throughout.

The workshops would be co-facilitated by the core evaluation team member and the National Consultant and note capturing will be shared to ensure views and perspectives are captured in the final report. Logistical support from the UNICEF CO will be requested where possible for each country case study.

Table 11: Proposed participants for in-country participatory workshops

In-country stakeholders	Purpose	Average no. of people/ country
UNICEF CO Staff	To understand the perceived importance of gender as part of individual roles and as a cross-cutting priority within their work and explore workplace culture in relation to gender identity and inclusion.	16 – 24

6.3.5 Top-up case study

To avoid evaluation fatigue in the ESARO region, the seventh case study will take a different format and consist of a ‘top-up’ to a recent evaluation undertaken by SDDirect for UNICEF – Formative Evaluation of Gender Transformative Programming through Investment in Adolescent Girls’ Leadership in Eastern and Southern Africa (2023). Mozambique, Namibia, Uganda, Malawi, Tanzania and Zimbabwe featured as case study COs. The 2023 evaluation applied a similar

²⁷ LILO approach : <https://evidenceforinclusion.org/wp-content/uploads/2021/08/LILO-Inclusion-Manual-May-2021.pdf>

methodology to the one put forward here: 1) adapted outcome harvesting against the Gender Equality and Social Inclusion (GESI) continuum and 2) participatory, youth friendly approaches. Outcomes were identified both through the document analysis as well as through stakeholder engagement (i.e., KIs and FGDs) and mapped against the GESI Continuum.

As a result, the current evaluation will seek to supplement the already collected data with further document review and interviews with UNICEF staff. The Evaluation Team will aim to work with the same National Consultant as the 2023 evaluation, to ensure comprehensive data collection and analysis for the 2024 evaluation; this National Consultant will receive support from a Core Evaluation team member.

6.3.6 Case study analysis and reporting

Data analysis will take place at country level to inform seven stand-alone CO case study reports. The Evaluation Team will use a **thematic analysis approach**²⁸ to analyse qualitative data from KIs, FGDs and workshop outcomes to identify key trends and patterns. Both inductive and deductive coding will be used against a base coding framework aligned to the evaluation matrix, including new thematic codes as trends are identified during analysis. All qualitative data will be assessed using Dedoose, a qualitative and mixed-methods data software. Triangulation across quantitative and qualitative findings will take place following the coding of the qualitative data. Where possible, sub-group analysis will take place to ensure that an intersectional lens is applied. All data collected during the case study will be disaggregated at minimum by gender, age and disability, and other characteristics as agreed in discussion with UNICEF and the YAG (e.g. sexual orientation, race, ethnicity) to ensure representation of a diverse range of voices and allow for sub-group analysis.

During the data analysis and reporting phase, a draft report will be produced in line with the UNICEF Evaluation case study template and GEROS standards. When the draft report is ready, a virtual validation workshop will be held with the CO to present the findings and draft recommendations. Each final CO report will be approximately 30 pages long, including findings, conclusions and CO-specific recommendations.

6.4 Data collected in an additional set of country offices

To complement the appreciative inquiry, in-person case studies, the evaluation team will conduct remote supplementary data collection activities with a series of country offices that did not achieve the 70% target for the GAP Standard Index Score.

The purpose of this data collection is to supplement the data collected through the in-depth case studies, to get a wider view of country office experiences of implementing the GAP, the relevance of the GAP and enabling and hindering factors. This will ensure that the viewpoints and positions of a wide range of COs are reflected in the evaluation and the scope of the challenges experienced and hampering factors are explored in greater depth.

The selection of these countries will take place during implementation, utilizing the data collected to select the case study country offices but applying new criteria. The team will select five countries that are not meeting the 'high gender performance' criteria used to select the case study countries. Specifically, we will select five countries that have not achieved the target for the GAP Standard Index Score (Target: 70%), a range of contextual criteria will be considered to ensure a range of country income types, office types, regions, national gender equality and norms etc. is represented.

It is proposed that a light-touch document review for these countries take place as part of the global document review. The evaluation team will conduct two remote group interviews with UNICEF staff in these country offices to explore the programmatic change pathway and institutional change pathway.

²⁸ Thematic analysis involves identifying and analysing patterns or themes within qualitative data. It is a flexible and adaptable approach that can be used to analyse data from various sources, such as interviews, focus groups, and open-ended survey responses.

Table 12: Proposed KIIs for in-country supplementary data collection

Stakeholders	Purpose	Average no. of people/ country
UNICEF CO Programmatic and Gender Staff (e.g. programme staff, MEL staff, Gender Focal point)	To explore the implementation of the GAP as relevant to programmatic change and goal level priorities.	one per country (up to seven people) (assuming a maximum of five country offices)
UNICEF CO Institutional Staff (e.g. HR and Finance)	To explore the implementation of the GAP as relevant to institutional or organizational culture shift.	one per country (up to seven people) (assuming a maximum of five country offices)

6.5 Sampling

This sampling table covers data collection for this evaluation by stakeholder type.

Table 13: Overall number of participants to KIIs, FGDs and workshops

	Inception	Global level (incl. comparative analysis)	Country-level (7 in-country case studies ²⁹ + interviews in up to 5 supplementary countries ³⁰)	
TOTAL per category, all countries				
UNICEF staff (HQ, RO, CO) (qualitative)	KIIs or FGDs with 15 people [15p]	KIIs or FGDs [25p]	2 group interviews (5 p.), or up to 10 KIIs with CO staff (10.p) *5 for top-up country [65 p.] Participatory workshop 16-24 p [140p.]	Approx. 158
UNICEF staff (HQ, RO, CO) - (quantitative)		All-staff survey [150]		Approx. 150
UN agencies/INGO / donors	-	KIIs [29p]	3 KIIs [21p.]	Approx. 50
Adolescent boys and girls	-	-	2 FGDs (5-8 p.) N/A for top-up country [78p.]	Approx. 78

²⁹ These are the countries that display positive results in integrating gender in systems and programmes, for which a case study report will be written.

³⁰ These are the additional countries, selected because of the challenges they experience, in which key informants interviews will be conducted remotely.

CSOs representatives (NGOs, networks, activists)/ academics/ other experts	-	-	4-6 KIIs (4-10 p.) [42.]	Approx. 42
Government decision-makers and policy makers	-	-	2 KIIs (2p.) N/A for top-up country [12.p]	Approx. 12
TOTAL	487			

6.6 Analysis and synthesis for the evaluation report

The data collected for this evaluation will be extensive and varied and this phase will involve the analysis across these different data sets and sources and within data sets. Analysis will be done to answer the evaluation questions and test the theory of change and assumptions and to understand the contribution of UNICEF to the outcomes articulated in the ToC. Analysis at this stage will include:

- **Analysis of trends** across HQ, Regional and Country level will be conducted to understand how the impact of the Gender Policy and the GAP is trickling down, to compare and contrast the results at each level and unpack the enabling and hindering factors.

Data sources: secondary data and document review, key informant interviews, case studies, survey.

- **Correlation analysis** will be undertaken at global level to explore the relationship between institutional enablers and programmatic results. This will be compared to trends identified through the case study country level data analysis.

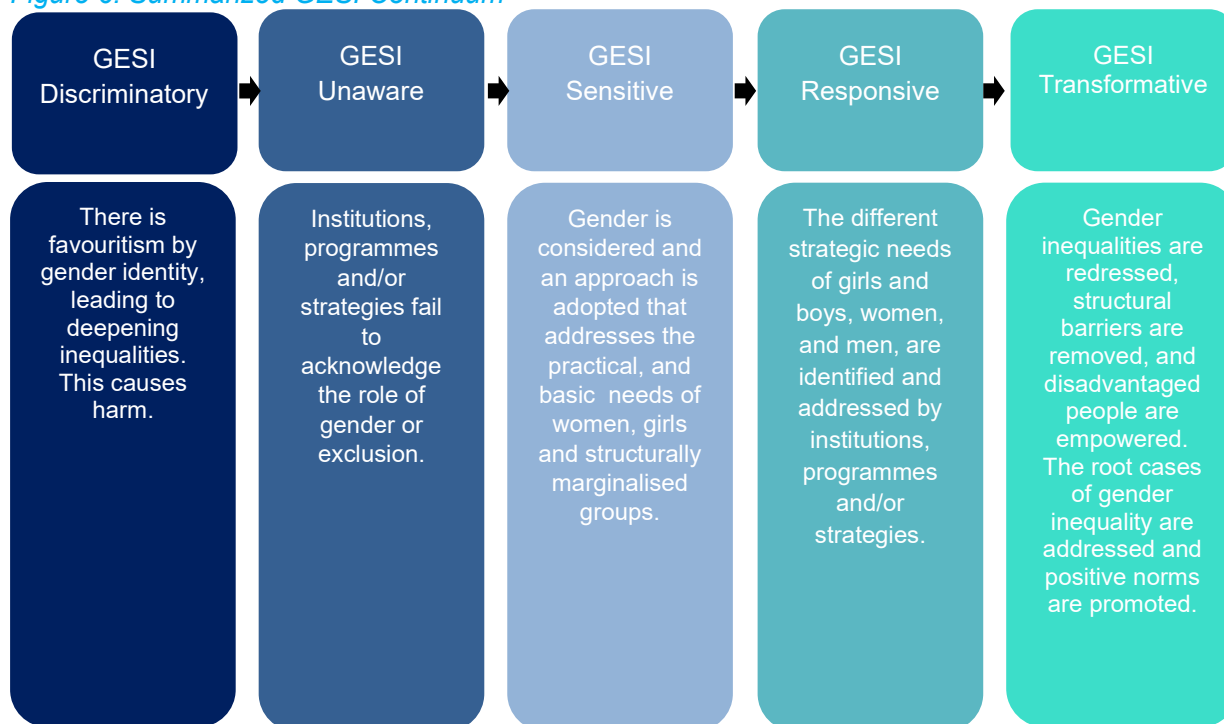
Data sources: secondary quantitative data.

- To draw trends from the case studies a **comparative analysis across case studies** will take place to examine emerging themes and areas of convergence and divergence across these different contexts with a particular focus on drawing out lessons on ‘what works’ to strengthen gender integration across different country contexts. This will be supplemented by the analysis of additional CO interviews to drawing out hindering factors and challenges.

Data sources: case studies and supplementary interviews.

- The **GESI continuum** will be used to form part of the analytical framework, the continuum presents a spectrum against which elements of UNICEF’s operations, processes, policies and programming (at global, regional and country level) can be understood and assessed, illustrating that achieving a transformative approach is a process, and will be influenced by various factors such as ambition and political will, resources, timeframe, capacity, mandate and contextual factors.

Figure 6: Summarized GESI Continuum



The **evaluation synthesis** will be a participatory process, the evaluation team will host a series of analysis and validation workshops throughout this phase:

- An **internal analysis workshop** will be conducted during the synthesis and analysis process, gathering all Evaluation Team members. The team will analyse and triangulate the findings of the various workstreams, and identify the headline findings under each research question, prior to drafting the evaluation report. To ensure an optimal use of triangulation and group work, the evaluation team proposes to conduct this workshop face to face.
- An **intermediary workshop** will be held following synthesis and analysis with the UNICEF Evaluation Office, ERG and YAG, where the team will present the preliminary findings and conclusions.
- A **recommendation co-creation workshop** is anticipated with the evaluation team, UNICEF evaluation office, the Gender Unit and other key stakeholders.

The evaluation report will fulfil the requirements of UNICEF standards for evaluation and UNICEF-Adapted UNEG Evaluation Reports Standards. The evaluation team are familiar with these as a result of previous evaluation work with UNICEF. As per SDDirect practice, the tools and deliverables produced throughout the data collection and analysis phase for this evaluation will be reviewed thoroughly by the Quality Assurance member of the team, ensuring they are GEROS compliant (see section 9).

7. DELIVERABLES

Over the course of the evaluation, the team will produce the following formal deliverables:

- One inception report;
- One short intermediary report (a descriptive analysis of GAP data and survey data);
- Six Cases study reports;
- One final evaluation report.

Table 14 below describes the deliverables and presentations that will be produced, with timelines.

Table 14: Key deliverables

Deliverable	Content	Date
Deliverable 1: Inception Report <i>Presentation of the draft to ERG and YAG</i>	Presents the scope of work and how it will be conducted.	14/03/2024 20/02/2024
Deliverable 2: Interim report <i>Presentation of findings: UNICEF EO, ERG</i>	Presents findings from a sub-set of our data analysis and collection: Descriptive analysis of the GAP standard index, and of the all-staff survey results (10-15 pages).	25/04/2024 06/05/2024
Deliverable 3a: Pilot Case study Report	It will respond to the EQ in the pilot country (DRC).	15/05/2024
Deliverable 3b: Case Study Reports	The report will present evaluation findings from each of the Country Offices visited (30 pages).	24/06/2024
Deliverable 4: Draft Evaluation report <i>Presentation of preliminary finding and co-creation of recommendations: UNICEF EO, ERG and YAG</i>	The draft evaluation report will present the methodology, findings, conclusions and recommendations of the evaluation (60 pages).	02/09/2024 12/08/2024
Deliverable 5: Final Evaluation Report <i>Presentation of findings</i>	The final report be prepared after one round of consolidated feedback from the client. This will be submitted in UNICEF branding, illustrated with data and infographics, and copy-edited	30/09/2024 16/09/2024
Deliverable 6: Dissemination products	Including: <i>Evaluation Summary Brief</i> 5-7 pages, professionally designed and youth friendly with infographics, in UNICEF branding.	30/09/2024

8. ASSUMPTIONS AND RISKS

Table 15: Risks and mitigation measures

Risks	Mitigation measures	Likelihood
The organizations selected for invitation to the comparative study do not wish to participate and we need to reduce the sample. It will take too much time to secure participants' commitment, set up meetings and convene for the final debrief workshop.	We have a list of priority organizations and possible back-ups. We will communicate clearly to potential candidates the benefits of participating and measures in place to safeguard confidentiality. We may adapt the method and the scope, depending on the time it takes to engage the participants.	Medium
Delayed access to raw data causes delays in the timeline for the evaluation and prevents our ability to sequence data collection.	Conversations with the Gender team MEL lead have been comprehensive. Relevant datasets and methods of obtaining them have been explored extensively in the inception period and plans for data sharing discussed.	Medium
Delays in collecting data in case study countries will delay finalization and submission of case study reports.	A firm deadline was set for data collection (end of May), to avoid a major knock-on effect on the subsequent reports. However, the submission time of the case study reports may be adjusted depending on team's member deployment and staff availability for internal QA	Medium
The availability of staff to participate in discussions on the draft report may be affected by the timing (during August holidays), which may hamper meaningful engagement.	We will ensure there is advance notice for the meeting and will remain flexible on the date.	Low
The extent to which we can comprehensively answer EQ 3 is in part determined by the quality, and reliability of programme and impact data collected by UNICEF CO programmes and the appropriateness of the M&E indicators in measuring results attributable to UNICEF interventions. There is a risk that if the quality of the data is poor or the outcome level indicators are not appropriate, this could impact the extent to which we can answer these questions.	The evaluation team is in close consultation with the Gender team on the selection of programme indicators and the country office evaluation and gender focal points to identify relevant country level data.	Medium

<p>The evaluation team is unable to visit one of the selected Country Offices for any of the following reasons:</p> <ul style="list-style-type: none"> • The CO does not wish to participate. • The CO cannot participate during the proposed data collection window. • The availability of the CO clashes with the availability of the core team. • The security situation worsens in the country between selection and visit. • This would impact methodology and timeline (reducing number of case studies completed). 	<p>A long list of countries has been selected with preferred and back up countries. Qualitative data on the motivation of country offices to participate has been considered in country selection. Worst case scenario, we would reduce the number of in-depth case studies conducted and top up with other sources and interviews.</p>	<p>Medium</p>
<p>The number of different languages, including local languages, presents logistical challenges.</p>	<p>National Consultants recruited who can work in local and national languages as well as English, French or Spanish, where appropriate.</p>	<p>Low</p>
<p>Some participants are not able to participate due to accessibility requirements.</p>	<p>We will enquire about any reasonable adjustments or accessibility requirements needed to ensure equitable and inclusive participation, including sign language interpreters and provision for carers and aides to accompany participants as required.</p>	<p>Medium</p>
<p>Challenges in recruiting national consultants delay the implementation of the case study data collection.</p>	<p>SDDirect has a wide network of consultants in many of the selected case study countries. As soon as we secure approval from UNICEF we will launch recruitment. We anticipate phasing the case studies to manage these different timelines. We will also be able to replace any national consultants should the need arise.</p>	<p>Low</p>
<p>Evaluation fatigue by key informants.</p>	<p>We have considered where evaluation visits have been conducted most recently to exclude those countries from our case study selection.</p>	<p>Medium</p>
<p>If the country offices operating in humanitarian contexts are unable to participate the sample will be skewed towards development contexts.</p>	<p>We have identified some back-up country offices operating in humanitarian contexts. If the sample becomes skewed, we might need to reach out to humanitarian contexts through the supplementary data collection.</p>	<p>Medium</p>

<p>Additional elements of methodology have been included during the inception phase, which require an increase in budget to deliver, this has been requested as part of the cost extension. If this is not granted we would be unable to deliver the proposed methodology in full.</p>	<p>Contract amendment has been submitted, additional days in the existing budget have been reallocated.</p>	<p>High</p>
<p>For quantitative data analysis, 2023 is not made available to the evaluation team as planned in February.</p>	<p>Preliminary report may only include data up to 2022, and depending on when this is made available, data for 2023 may be included only in the final evaluation report.</p>	<p>High</p>

Table 16: Assumptions and mitigation measures

Assumptions	Mitigation measures
<p>Selected countries are/remain safe/feasible to work in.</p>	<p>If needed, we will draw on a reserve list of countries from the long list of all countries in the region, and if required we will reduce the number of case study countries.</p>
<p>UNICEF and other UN staff will be willing and have time to engage in the evaluation.</p>	<p>Evaluation team is careful and proportional about requests for time and support.</p>
<p>Documentation is readily available for the evaluation team.</p>	<p>Working closely with the UNICEF Evaluation Office, Gender team and Regional Offices and Country Offices to ensure documentation is available in a timely and accessible manner.</p>
<p>UNICEF country team and assessment team able to mobilize for the pilot country visit within 5 weeks of submission of inception report.</p>	<p>Planning for pilot country visit to begin as soon as IR is submitted, to be led by the TL who is holding time for the pilot window and remaining flexible on destination.</p>
<p>Piloting the use of an AI tool to explore a large volume of qualitative data, will yield the expected results.</p>	<p>The review of the qualitative documentation will revert to traditional methods, aligned to human capacities and to a sample of documentation (extensive document review in the countries selected and a sample of documents produced at global and regional levels).</p>

9. EVALUATION GOVERNANCE, MANAGEMENT ARRANGEMENTS AND

Figure 7: Team diagram

EVALUATION TEAM



Evaluation Governance and Management Arrangements

Erica Mattellone, Senior Evaluation Specialist and Institutional Effectiveness Team Lead, is the evaluation manager for this evaluation. She has overall approving authority for the project, and will manage the exercise to ensure impartiality, independence and credibility of the work. She will work in coordination with the Gender team responsible for coordinating the GAP, Programme Group and other Headquarters Divisions, Regional Offices and Country Offices, who will be responsible for ensuring access to information and key stakeholders to ensure that the evaluation will produce relevant and reliable findings and actionable recommendations.

Barsha Pradhan, Evaluation Officer, has been embedded in the evaluation team to provide access to information and key stakeholders, and will support the coordination of the data collection and the participation of the Youth Advisory Group members.

Evaluation team members roles and responsibilities

Celine Mazars, Team Leader, will provide a technical oversight and management role, coordinating and supervising the work of the evaluation team and liaising with the UNICEF evaluation manager and key stakeholders to ensure coherence across the project. She will be responsible for ensuring the quality of the process from start to finish, outputs, methodology and timely delivery of all products, including adherence to ethical standards adherence during all phases of the evaluation (with support and collaboration of the DTL). Celine will lead the first case study, piloting the approach and tools and will provide technical oversight for the remaining case studies and data collection.

Alix Clark, Deputy Team Leader, will work closely with the Team Lead on all the deliverables of the assignment. Responsibilities include co-leading the design of the evaluation and initial desk review, analysis and reconstruction of theory of transformation, formulation of evaluation questions, co-writing the inception report, producing draft findings in the form of intermediate products and co-writing the final evaluation report. The deputy team lead will take on leadership of facilitation/engagement with the Youth Advisory Group and design of the quantitative survey elements. Alix will have responsibilities for leading specific case studies and supporting the Team

Lead in oversight.

Florianne Gaillardin, Senior Gender Consultant, will support the Team Lead in the conduct of the evaluation, ensuring the application of gender equality and feminist evaluation principles and will work closely with the DTL to contribute to all deliverables through data collection, analysis, presentation and report writing. Florianne will play a significant role in the assessment for institutional GAP enablers and lead the design and delivery of the benchmarking exercise, working closely with the DTL, Florianne will also support the researchers in their roles.

Sonal Zaveri, Senior Gender Consultant, will support the Team Lead in the conduct of the evaluation, ensuring the application of gender equality and feminist evaluation principles. Sonal will contribute to all deliverables through data collection, analysis, presentation and report writing. Sonal will take a lead role in the case study deliverable for this evaluation, working closely with the TL and DTL to oversee the implementation of this deliverable. Finalizing data collection tools, leading the rollout of training for the National Consultants in preparation for the case studies (including ethical and safeguarding training), completing one case study and providing support to the evaluation team. Sonal will also contribute to analysis and triangulation efforts, and take a lead role in the final evaluation report.

Lucia Soldà, Researcher, will be supporting the team in data collection, analysis and developing key products. This will include support to the initial desk review, analysis and reconstruction of theory of transformation and support to survey design. She will also conduct desk review, in-depth document reviews and data collection in the form of KIIs (internal and external), FGDs and workshops for case studies, and support the analysis of findings and the write-up of the evaluation report. Lucia will be involved in the case studies and will be paired with a Senior National Consultant who will lead (supported by the senior gender consultant).

Hanna Smit, Researcher, will be supporting the team in data collection, analysis and developing key products. This will include support to desk/document reviews analysis and reconstruction of theory of transformation. She will also conduct desk review, in-depth document reviews and data collection in the form of KIIs (internal and external), FGDs and workshops for case studies, and support the analysis of findings and the write-up of the evaluation report. Hanna will be involved in case studies and will be paired with a Senior national consultant who will lead (supported by the senior gender consultant).

Jo Feather, Quality Assurer, will provide quality assurance (QA) on all deliverables and ensure that all deliverables are GEROS compliant. She also takes institutional accountability for delivery.

Isadora Brizolara, Project Manager. Provides overall project and risk management support, including client liaison and management, contracting and budgeting for national consultants.

National Consultants, conduct detailed stakeholder mapping of stakeholders in country, in consultation with UNICEF CO and TL. They will lead on securing ethical clearance with relevant national bodies; conduct interviews and focus groups with UN, government and implementing partners, and young people; support qualitative data collection, support the sensemaking / analysis by the young people; and co-produce country case study reports with annexes.

Quality assurance

We apply technical quality assurance (QA) to all project deliverables, in accordance with our internal company QA policy. This requires us to integrate QA into all stages of the project cycle, using peer review across the core Evaluation Team, and our senior in-house experts to QA all written outputs.

A Principal Consultant is responsible for ensuring that the design and structure of all projects meet quality assurance and ethical standards. For this assignment, Jo Feather will be responsible for QA on relevant outputs. The Team Leader, Celine Mazars, will have responsibility for overall delivery, project management and quality assurance as the project progresses, liaising with the

UNICEF focal point, on a regular basis.

Our diverse, multi-disciplinary team offers exceptional knowledge and expertise in gender equality and women's empowerment, feminist evaluation approaches, mixed-methods, theory-based, utilization-focused, participatory, and gender and human rights focused evaluations who alongside our team members from the region and case study countries provide a deep understanding of the local context and necessary language skills.

Clear roles and responsibilities have been assigned to specific team members with clearly defined lines of accountability between team members and UNICEF to support quality deliverables. We will ensure regular and transparent communication and coordination between all team members, as a key risk management strategy, in allowing the team to identify, communicate and mitigate any risks at the earliest opportunity. Clear and regular dialogue with UNICEF will ensure that requirements and expectations are agreed and understood throughout the evaluation, and relevant information is shared throughout the evaluation (e.g., earlier deliverable drafts for comments if necessary).

Finally, technical inputs from ERG and YAG members will also serve to improve the quality of the deliverables produced by the evaluation team.

10. WORKPLAN

	February				March				April				May				June				July				August				September								
	05/02/24	12/02/24	19/02/24	26/02/24	04/03/24	11/03/24	18/03/24	25/03/24	01/04/24	08/04/24	15/04/24	22/04/24	29/04/24	06/05/24	13/05/24	20/05/24	27/05/24	03/06/24	10/06/24	17/06/24	24/06/24	01/07/24	08/07/24	15/07/24	22/07/24	29/07/24	05/08/24	12/08/24	19/08/24	26/08/24	02/09/24	09/09/24	16/09/24	23/09/24	30/09/24		
Activities																																					
Global and Regional Data collection and analysis phase																																					
All staff survey																																					
Qualitative tool development																																					
Analysis of Secondary Data																																					
Global Document review																																					
Staff interviews and FGDs, and analysis																																					
Comparative Study Data collection and Document Review																																					
Deliverable 2: Interim report																																					
Presentation of findings																																					
Action feedback on Interim Report																																					
Case Study preparation and pilot case study																																					
Tool development																																					
Case Study team training																																					
Country level secondary data analysis																																					
Pilot case study: data collection and analysis																																					
Deliverable 3.1 Pilot Case Study Report																																					
Address Case study Feedback from Client (one Consolidated round of feedback)																																					
Case studies																																					
Kick Off Calls																																					
Country Level Document and secondary data review																																					
Case study Data collections and coding.																																					
Deliverable 3.2 Case Study Reports submission																																					
Address Case study Feedback from Client (one Consolidated round of feedback)																																					
Supplementary data collection																																					
Draft evaluation report																																					
Internal analysis workshop																																					
Synthesis of findings																																					
Preliminary findings workshops with HQ, ERG and YAG																																					
Deliverable 4: Draft Evaluation Report																																					
Final report, presentation and other communication products																																					
Findings validation presentation																																					
Deliverable 5: Final report																																					
Communication/dissemination products																																					