

EXTERNAL INDEPENDENT ASSESSMENT OF THE IMPLEMENTATION OF UN WOMEN'S EVALUATION POLICY

November 2025

Dr Tom Ling

Dr Ananda S. Millard

Introduction to this Report

This assignment examines the implementation of UN Women's evaluation policy between 2020 and 2025. This report summarizes the results of an independent assessment carried out between September 2025 and November 2025. The Terms of Reference (ToR) state:

The assessment will take into consideration contextual and organizational changes, as well as ongoing UN and UN Women change management processes. The assessment will focus on IES-led independent corporate evaluations and decentralized evaluations, including systems for evaluation planning, oversight and quality assurance. The coordination role of UN Women on gender-responsive evaluation in the United Nations system and national evaluation capacity development workstreams will also be included in the assessment.

More specifically this assignment considers three overarching questions:

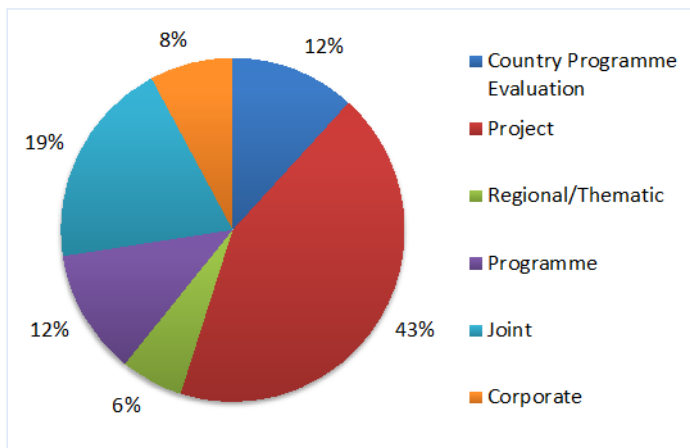
1. **Relevance of the policy:** *the extent to which the evaluation policy meets professionally recognized international standards, the needs of UN Women stakeholders, and benchmarking with relevant comparators. Also, the extent to which the key performance indicators managed outside the policy provide appropriate and timely information on UN Women's evaluation function.*
2. **Independence, credibility and utility:** *the degree to which the policy provides safeguards for independence, and the extent to which the credibility and utility of evaluations have been maintained and supported. Also, examine implementation of quality assurance systems for evaluations, such as the Global Evaluation Reports Assessment and Analysis System, for timely and robust evaluation evidence.*
3. **Coverage:** *the extent to which the criteria defined in the policy and the IES coverage norms ... provide adequate and appropriate coverage in support of the effective, efficient and impactful implementation of UN Women's mandate and strategic plan, SDG 5, as well as the growing demand for development effectiveness and efficiency measures both within UN Women and among its partners. Also, identify any challenges affecting the overall implementation of the policy.*

These questions are fully addressed in this Report. In addition, given the scale and speed of changes in the external context of UN Women (including developments in the evaluation community), we consider what this might imply for the continuing delivery of the evaluation function of UN Women. Accordingly, *'The primary purpose of this assessment has been to identify key achievements and challenges in the implementation of the policy and provide insights to further strengthen implementation efforts.'*

Context and background

The United Nations Entity for Gender Equality and the Empowerment of Women (UN Women) is dedicated to the achievement of gender equality and the empowerment of women. The evaluation policy is intended to ensure that evaluations support these aims through strengthening accountability and learning. To this end, the successful implementation of the evaluation policy would include the systematic adoption of high professional standards and the use of evaluations to learn and achieve better outcomes for women and girls. This assessment considers both criteria across the various types of evaluation. The distribution of evaluations by type is shown in Figure 1. In total there were 51 reports assessed in the 2023-24 GERAAS assessment.

Figure 1 Distribution of evaluations by type.



Source: GERAAS Assessment April 2023-Jan 2024

As part of our assessment, we consider the role played by different parts of the evaluation system, including the Executive Director, the Executive Board, the Independent Evaluation Service (IES), the Global Evaluation Reports Assessment and Analysis System (GERAAS), and organizational management and leadership. IES’s key responsibilities include conducting corporate and strategic evaluations, providing technical assistance to decentralised evaluations, ensuring quality assurance for all evaluations, and advising on UN Women’s strategies and operations. It also contributes to system-wide evaluations and develops methodologies to promote gender-responsive evaluation practices across the UN system. It acts as the custodian of evaluation standards and promotes the role of evaluation within UN Women.

As discussed below, the implementation of the evaluation policy of UN Women adheres to the professional standards established by the United Nations Evaluation Group (UNEG), especially evident at the corporate level, and continues to perform well against the prescribed KPIs across all evaluations (Finding 1). These are commendable. However, there are caveats relevant to each of these achievements.

Professional standards in the evaluation field continue to evolve.¹ There is no single measure of industry standards, but in recent years there have been important developments relevant to implementing the evaluation policy at UN Women. On the demand side, there is a growing volume of alternative sources of evidence for policy makers and an increased questioning of

¹ The Future of Evaluation is discussed in a World Bank IEG YouTube from April 2024 at <https://www.youtube.com/watch?v=CyglGIFUOs&t=71s> See also: <https://europeanevaluation.org/wp-content/uploads/2020/03/ees-newsletter-2018-05-may-r12.pdf>

expert opinion through social media and competing claims around 'fake news'; UN Women is not insulated from this problem. Possibly, even more significantly, are the changes on the supply side – growth of new sources of data that will most likely transform how the business of evaluation is conducted: A.I., digital, and citizen-generated data. Methodologically, we are seeing important shifts, including the rise of participatory evaluation, real time evaluation, continuous learning for agile management (sometimes called 'evaluative thinking'), benchmarking and comparable datasets, and calls for more ex ante rather than ex post evaluations. The implementation of the evaluation policy must be able to respond to these opportunities and challenges.

Also, meeting KPIs is an important indicator of a well-managed evaluation system, but it is not in itself proof that evaluation in UN Women is contributing to the empowerment and equality of women and girls. KPIs in evaluation management are part of a wider institutionalization of evaluation systems in large organizations² and there is some ambivalence about whether this has resulted in the increased usefulness of evaluation. KPIs on their own are unable to provide context, require interpretation, may suffer from selection bias or definitional weakness, and may miss qualitative aspects.³ While institutionalizing evaluation provides an important way to promote accountability and evidence-informed decision making, on its own it is insufficient and it may become bureaucratized, politicized, or ritualized.⁴

The modalities of achieving change for women and girls are also changing in ways that matter for how evaluation is managed. New ways of delivering SDGs will increasingly require evaluators to examine cross-UN initiatives and private/public/NGO collaboration and the integration of sustainability indicators into M&E frameworks. The most recent MOPAN Report on UN Women states: 'One of the most pressing issues in strategic management is the difficulty in translating high-level policy influence into sustained, systemic change, particularly at the country level.'⁵ Even more recently, the MOPAN report 'Closing the UN Mandate-Results Gap: Realigning Incentives for Lasting Impact'⁶ emphasises that given the changing role of UN entities:

the difficulties of articulating, monitoring, and implementing integration between functions... should "make maximum use of comparative advantages" and "facilitate constant two-way feedback." In practice, however, no system is in place to monitor, evaluate, and learn from these linkages.

More widely, MOPAN emphasise the need to institutionalise learning into policy. Evaluative thinking (see below) should play a central role in this.

² For example, we have seen a significant institutionalization of evaluation within the institutions of the European Union: Stephenson, P. J., & Schoenefeld, J. J. (2023). The role and functioning of evaluation in the European Union. In F. Varone, S. Jacob, & P. Bundi (Eds.), *Handbook of Public Policy Evaluation* (pp. 266-284). Edward Elgar Publishing. <https://doi.org/10.4337/978180884892.0026>

³ <https://www.evalcommunity.com/career-center/key-performance-indicators/>

⁴ As Cordoncillo notes, 'However, the institutionalisation of evaluation is certainly necessary, but it is not sufficient to guarantee the use of evaluation nor to consolidate an evaluation culture. In this sense, it is especially worrying that none of the mechanisms identified in the scoping review are aimed at promoting the values that are usually associated with the culture of evaluation. Without solid values, institutionalisation runs the risk of becoming mere bureaucracy.' Carla Cordoncillo Acosta, *Mechanisms for institutionalising evaluation: A scoping review*, *Evaluation and Program Planning*, Volume 114, 2026, 102710, ISSN 0149-7189,

⁵ MOPAN (2025), *MOPAN Assessment Report: UN Women (Part I)*, MOPAN, Paris © MOPAN 2025.

⁶ MOPAN 2025 *Closing the UN Mandate-Results Gap: Realigning Incentives for Lasting Impact* <https://www.mopan.org/en/our-work/performance-insights/multilateral-effectiveness-in-a-shifting-landscape/mandates-thematic-brief-summary/full-report.html>

In the following sections we identify 10 key findings and 6 opportunities for further progress. This is in the context of a well-managed and maturing evaluation system, in which, never-the-less, there are some persisting problems and emerging challenges.

Finding 1. There has been continuing progress towards meeting KPIs

Overall, the implementation of the evaluation policy within UN Women has continued to mature since UN Women was established in 2010. Table 1 shows steady and commendable progress towards meeting its KPIs.

Table 1 Performance of the Evaluation Function in UN Women Key Performance Indicators

Key Performance Indicator	Description	2018	2019	2020	2021	2022	2023	2024	Target %
Financial resources invested in evaluation function	Evaluation expenditure over UN-Women programme expenditure	2.2	2.0	1.9	1.8	1.9	2.3	1.6	2-3
Human resources for monitoring and evaluation	Offices that appointed an M&E focal point or officer	95	98	98	98	98	97	99	100
Evaluation coverage	At least one evaluation per Strategic Note cycle (2017- 2021)	84	86	92	97	93	98	96	100
Evaluation implementation rate	Percentage of evaluations being implemented (completed and initiated)	90	77	95	97	92	84	87	85
Quality of evaluation report	Percentage of evaluation reports rated “Good and above”	88	68	90	98	86	100	100	100
Joint evaluations	Percentage of joint evaluations contributed by UN-Women	N/A	N/A	N/A	N/A	29	27	35	25
Management response submission to GATE	Percentage of completed evaluation reports submitted with management response to GATE	97	95	95	97	100	100	100	100
Implementation of management response	Percentage of management response key actions being implemented	87	85	89	85	91	92	86	80 (95 by 2025)
Use of evaluations	Percentage of offices that reported using evaluation	86	86	90	94	87	89	95	90 (95 by 2025)

Sources: Compiled from *Report on the Evaluation Function of the United Nations Entity for Gender Equality and the Empowerment of Women 2023*, *Report on the Evaluation Function of the United Nations Entity for Gender Equality, the Empowerment of Women 2021 Annual Report*, and *Report of the evaluation function of the United Nations Entity for Gender Equality and the Empowerment of Women 2024*.

The Director of the Independent Evaluation, Audit and Investigation Services reported that this solid improvement in delivery against expectations has been consistent up to the time this Note was written. In the year up to September 2025⁷ the IES completed three corporate evaluations, a synthesis of UN Women’s performance against the strategic plan, alongside the synthesis of country portfolio evaluation that was in progress. The Director reported that fifty evaluations were conducted, or were in progress, at the decentralized level.⁸ The IES is responsible for overseeing and coordinating GERAAS, and provides guidance to support the conduct of decentralized evaluations, but it does not directly manage decentralized evaluations. This is done at the regional or country level. IES commissions external independent consultants to assess and rate evaluation reports using UNEG criteria to create the scores to be found in Table

⁷ Executive Board of the United Nations Entity for Gender Equality and the Empowerment of Women Second regular session of 2025 9–10 September 2025 Item 1 of the provisional agenda Organizational matters

⁸ However, we note that financial expenditure on evaluation in 2024 fell below the 2 per cent target and the implementation of management responses fell to 86 per cent following the introduction of a more rigorous monitoring and tracking system.

1. These scores are compiled and included in its annual report to the Executive Board. Final evaluation reports, along with their GERAAS rating, are available through the Global Accountability and Tracking of Evaluations (GATE).⁹

While the GERAAS system and scores are commendable, they are indicators of operational effectiveness rather than necessarily reflecting how far evaluation contributes to the broader objectives of UN Women. The KPIs used generally capture outputs such as the number of evaluations completed or systems improved, rather than the more complex and long-term outcomes associated with gender equality and the empowerment of women, which lie at the core of UN Women's objectives. Therefore, while meeting the targets identified in Table 1 demonstrates operational effectiveness, these KPIs do not automatically show that evaluations are informing strategic decisions or contributing to meaningful change on the ground.

The approach taken by UN Women with regard to KPIs is similar to the combination of internal review and external oversight that are found in other UN Agencies. This involves setting standards based on UNEG criteria, engaging independent quality assessors working with standardized templates to allow comparison across geography, policy fields, and across time. The World Food Programme also engages with the Active Learning Network for Accountability and Performance (ALNAP). The UN Office on Drugs and Crime adopts a very similar approach with independent consultants working to a Quality Assurance template. Reporting of quality is very similar to the approach found in twelve other UN Entities. In the most recent cross UN data we could find (2017)¹⁰ all but two agencies used ordinal scales of measurement, and all used between 4-6 response categories. There was more variation over the measure of central tendency (medians, means, qualitative etc) but this is not core to the system. Entities varied equally over whether or not to weight sections. GERAAS was highly regarded in this review, noting that the written justifications for scores was a strength as well as using a qualitative aggregation for an overall score.

UN Women operates a quality assurance system that is in line with common UN practice (based on UNEG criteria), and it can demonstrate a generally steady or improving performance on KPIs. We can conclude both that that IES is overseeing an effective quality assurance system and that the results indicate that evaluations mostly meet these quality assurance criteria.

Finding 2. Governance arrangements and planning processes are understood and transparent but require further consideration

The UN Women¹¹ Strategic Plan 2022-2025¹² articulates the organisation's strategy, and in turn the UN Women Evaluation strategy supports the Strategic Plan in four key areas of work, as identified in the UN Women-IES Global evaluation strategy 2022-2025:¹³

- i. Implementing strategic corporate, regional and country portfolio evaluations;
- ii. Implementing effective decentralized evaluation systems;

⁹ <https://gate.unwomen.org/>

¹⁰ Iker Llabres December 2017 Evaluation quality assessment methodology in the UN system and changes to the ILO's quality appraisal methodology ILO Evaluation Office

¹¹ The Executive Board of UN Women operates within the governance structures of the UN, consists of 41 member states, provides operational guidance to UN Women, and reports annually to the UN General Assembly.

¹² <https://www.unwomen.org/en/un-women-strategic-plan-2022-2025>

¹³ <https://www.unwomen.org/sites/default/files/2022-06/UN-Women-IES-Global-evaluation-strategy-2022-2025-en.pdf>

- iii. Supporting the UN and national stakeholder partnerships on gender-responsive evaluations; and
- iv. Strengthening evaluation use.

The purpose of the evaluation policy document published in 2020 was to establish the context for that policy, evaluation principles and procedures; roles and responsibilities; quality assurance principles; the anticipated contributions to system-wide evaluation for gender equality and empowerment of women¹⁴. The 4-yearly corporate evaluation plan aims to align the evaluation portfolio with UN Women's organisational priorities and budgets.

The Executive Board endorses and reviews the evaluation policy and is informed of corporate evaluation plans and corporate evaluations, and may request corporate evaluations. The Executive Director is the champion of evaluation and is responsible for protecting the independence of the IES, and ensuring that it is adequately staffed. The IES is the custodian of the evaluation function, reporting directly to the Executive Director and its main responsibilities are to:

- a) establish and implement effective corporate evaluation systems
- b) strengthen decentralized evaluation systems
- c) lead UN coordination on gender-responsive evaluation
- d) promote innovation and knowledge generation concerning what works in promoting gender equality, through national evaluation capacity building.
- e) work to ensure that evaluation findings are used in the governance processes of UN-Women and are presented to and used by programme managers in the conduct of their work.¹⁵

The Head/Director of IEAIS and the Chief of Evaluation lead UN coordination on gender responsive evaluations, and support the implementation of the IES workplan, while promoting UNEG norms and standards. Regional Evaluation Specialists report functionally to the Chief of Evaluation and are located in regional offices. They conduct independent evaluations and manage strategic decentralized evaluations at regional and country levels. Regional Directors play an oversight role with respect to decentralized evaluations in their region, and are responsible for ensuring adequate staffing and competencies for the fulfilment of evaluation roles, including the mandatory recruitment of monitoring and evaluation specialists or appointment of monitoring and evaluation focal points.¹⁶ As noted above, the IES is responsible for overseeing GERAAS and presenting the results annually to the Board.

As well as overseeing GERAAS and being the custodian for UNEG principles, the IES leadership has strategic conversations, drawing on evaluation findings, with the Executive Board. In 2024, the Advisory Committee on Oversight (ACO) welcomed IES's focus on leveraging evaluation findings to inform corporate planning and decision-making and singled out the good collaboration between IES and the Strategic Planning Unit and IES's synthesis of UN Women's performance against the strategic plan supporting strategic alignment still further.¹⁷

¹⁴ <https://www.unwomen.org/en/digital-library/publications/2020/08/evaluation-policy-of-the-united-nations-entity-for-gender-equality-and-the-empowerment-of-women>

¹⁵ <https://www.unwomen.org/en/digital-library/publications/2020/08/evaluation-policy-of-the-united-nations-entity-for-gender-equality-and-the-empowerment-of-women>

¹⁶ <https://www.unwomen.org/en/digital-library/publications/2020/08/evaluation-policy-of-the-united-nations-entity-for-gender-equality-and-the-empowerment-of-women>

¹⁷ 'Report of the Advisory Committee on Oversight for the period from 1 January to 31 December 2024' June 2025

At a time when the practice of evaluation within UN Women needs to respond flexibly to a changing environment (as we argue below), and at a time when the tasks and priorities of UN agencies may be shifting, arrangements will need to be agile to ensure that the actions required to maintain the relevance and utility of the evaluation system are taken up at the most senior levels at speed. As the requirements for successful implementation of the evaluation policy change, senior engagement might be required to, for example, deliver improvements to data quality and relevance, to strengthen the links between evaluative thinking and programme delivery, to connect learning from programmes to advice to stakeholders, to ensure rights holders have an active voice in evaluation at the decentralized level, and to provide data that can deliver a clear account of the impacts achieved by UN Women.

The Mopan 2025 review,¹⁸ while noting significant improvements in the evaluation function (among other things), highlighted the need for decisions that could only be fully dealt with by senior decision-makers (emphasis added):

*Despite improvements in RBM, evaluation culture, and performance tracking tools, persistent gaps exist in data quality, timeliness, and field-level monitoring. Limited evaluation coverage, selection biases in case studies, and recurring recommendations on the same issues in evaluations, thus raising concerns about whether lessons are systematically prioritised and acted upon, diminish UN Women's ability to adapt programming effectively. **Decision-makers should prioritise investments in evaluation capacity, streamline reporting requirements, and enhance accountability for integrating findings into strategic planning.***

The implementation recommendations identified in the MOPAN Report will require executive stakeholders to be involved, along with the active engagement of the IES leadership. This resonates with the Report of the Advisory Committee on Oversight for the period 1 January to 31 December 2024 which emphasised the scale of the challenge facing UN Women and states: 'A strong senior leadership and well-co-ordinated roadmap with clearly defined milestones are crucial to ensuring alignment across the multiple demands of these initiatives, while maintaining a robust corporate governance framework, resource efficiency, and effective control mechanisms.'¹⁹ The evaluation function is a core part of this ambition.

Finding 3. Professional evaluation standards within UN Women are strengthened by the two-way flow of engagement with the wider UN system

With particular relevance to upholding professional standards in implementing evaluation policy, UNEG provides a transparent framework, ensuring alignment with established norms and practices, particularly in relation to independence, quality and utility of the evaluation system within UN Women, and its contribution to UN system wide gender-responsive evaluation.²⁰ This function of UNEG principles is common across UN agencies (albeit with nuances in how it is implemented). These standards are reinforced further through UN Women's multiple partnerships with other UN agencies, including in inter-agency humanitarian evaluations and UN system-wide evaluations and initiatives with the intention that evaluations

¹⁸ MOPAN ASSESSMENT REPORT United Nations Entity for Gender Equality and the Empowerment of Women (UN Women) Draft Assessment Report, Part 1 https://www.unwomen.org/sites/default/files/2025-04/embargoed_mopan_institutional_assessment_report_-_un_women.pdf

¹⁹ Report of the Advisory Committee on Oversight for the period from 1 January to 31 December 2024

[file:///C:/Users/tting/Downloads/UNW_2025_3_Add.1-EN%20\(1\).pdf](file:///C:/Users/tting/Downloads/UNW_2025_3_Add.1-EN%20(1).pdf)

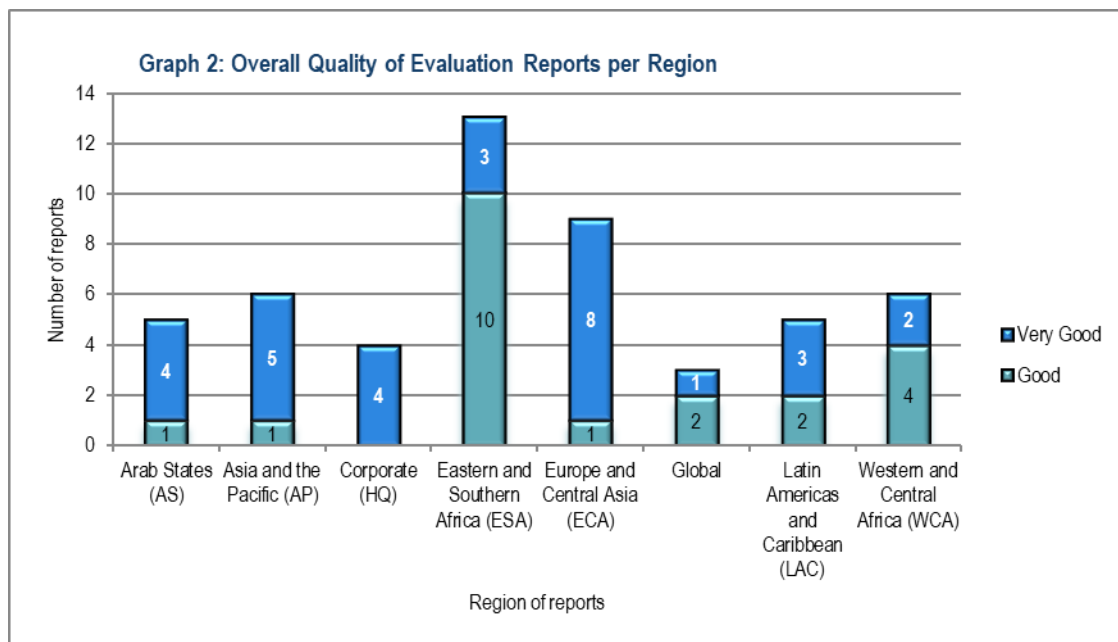
²⁰ [https://www.unevaluation.org/agency_profile/united-nations-entity-gender-equality-and-empowerment-women#:~:text=While%20aligned%20with%20the%20United,of%20Evaluation%20\(GATE\)%20system.](https://www.unevaluation.org/agency_profile/united-nations-entity-gender-equality-and-empowerment-women#:~:text=While%20aligned%20with%20the%20United,of%20Evaluation%20(GATE)%20system.)

incorporate gender equality. One important illustration of how working with other UN agencies can shape evaluation practice, and the subject of a corporate evaluation, is UN Women’s support for women’s participation in the peace process.²¹ Recent evidence strongly suggests that this role is often vital to peace and security but that further evaluations are needed to understand how, and under what circumstances, such efforts are most likely to succeed (not least with a view to strengthening UN Women’s policy influence in this area).²²

UNEG norms and standards underpin discussions of good practice within the UN Women evaluation system and provide a framework for maintaining professional standards. Organisations can self-assess their performance against a maturity matrix²³ which involves a clear rubric providing descriptors of the standards required to achieve performance levels ranging from 1-4 across the expected performance areas. UN Women also played a leading role in developing the UNEG handbook on gender and human rights which reinforces good practice within UN Women as well as promoting the visibility and importance of gender responsive evaluation across UN agencies.

The most recent independent assessment of the GERAAS results suggests high and improving adherence to UNEG standards, with some variation across regions and with corporate evaluations, as shown in Figure 2.

Figure 2 Quality of Reports by Region.



Source: GERAAS Assessment April 2023-Jan 2024

²¹ Report on the corporate evaluation of UN-Women’s support to women’s participation in peace processes https://www.unwomen.org/sites/default/files/2025-05/report_on_the_corporate_evaluation_womens_participation_in_peace_unw.2025.crp_.3_en.pdf

²² UN Women (2025) Facts and figures: Women, peace, and security 20 October 2025 <https://www.unwomen.org/en/articles/facts-and-figures/facts-and-figures-women-peace-and-security>; European Parliament 2025 Women in peace building: why and how to increase their role <https://www.europarl.europa.eu/topics/en/article/20250227STO27080/women-in-peace-building-why-and-how-to-increase-their-role>; Council on Foreign Relations undated Women’s Participation in Peace Processes <https://www.cfr.org/womens-participation-in-peace-processes/index.html>

²³ UNEG Reference Document 2022 UNEG Maturity Matrix Checklist Working Group, Peer Review Working Group https://www.unevaluation.org/uneg_publications/uneg-maturity-matrix-checklist

It should be noted that UN Women operate with a four-fold rating scale. 'Very good' is defined as 'The report can be used with high level of confidence and is considered a good example' and 'Good' as 'The report can be used with certain degree of confidence'. 'Fair' is 'Partially meets requirements with some missing elements. The report can be used with caution', and Unsatisfactory is 'Misses out the minimum quality standards.' In our view these descriptors encourage the use of 'very good' and 'good' while 'unsatisfactory' sets a very low bar for inclusion. UN Women should consider moving to a six-fold rating scale to allow for more nuanced reporting (as well as rewarding what is outstanding or excellent work). Be that as it may, the independent assessment shows (using consistent criteria over time) that reported results generally improved from 2019 to 2023, further underpinning a sense of a well-managed system.

However, in our interviews, concerns were expressed about how well these numbers captured the quality and consistency of evaluation capacity at the decentralized level, particularly regarding the availability of qualified human resources and the depth of professional judgement. It was also felt that the level of independence of the evaluation function varied by region, influenced by senior leadership, and this perception is partly reinforced by the results shown in Graph 2. We recognise that IES has the authority to influence but not to deliver these decentralized operations.

Finding 4. External reviews reinforce the sense of 'steady progress' in the evaluation system and IES is seen to respond positively to recommendations

UN Women's evaluation system is subject to continuing external scrutiny. External reviews are broadly positive and where improvements are identified IES shows a largely responsive engagement, as evidenced in reports reviewing responses to past recommendations and interviewees' comments. External reports include MOPAN, the Advisory Committee on Oversight (ACO), and external assessments of evaluation policy (such as this Report and the 2024 GERAAS assessment referenced above).

The ACO assists the Executive Director in fulfilling her oversight responsibilities in accordance with relevant best practices.²⁴ At least one interviewee observed that senior management were responsive to both formal and informal feedback from external scrutiny and that this had matured over time. Indeed, the 2025 MOPAN UN Women Report mentioned above notes that 'UN Women has demonstrated notable progress across multiple dimensions of its operations since the last assessment' (p.13). The Report of the evaluation function of the United Nations Entity for Gender Equality and Empowerment of Women 2023²⁵ noted that UN Women's key performance indicators remained 'on track' (see Table 1). The general adherence to standards and performance against KPIs, as reported here, resonates with other external reviews in the past 5 years or so.

²⁴<https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/About%20Us/Accountability/UN-Women-Advisory-Committee-on-Oversight-Terms-of-reference-en.pdf>

²⁵ Report of the evaluation function of the United Nations Entity for Gender Equality and Empowerment of Women 2023 UNW/2024/4

Finding 5. There is a systematic approach to conducting corporate and decentralized evaluations, supported by comprehensive formal and accessible quality assurance mechanisms, but decentralized evaluations require further consideration

The Corporate Evaluation Plan (CEP) provides a framework for the systematic collection of evaluative evidence covering the DAC-OECD criteria of relevance, coherence, effectiveness, efficiency, impact and sustainability of work completed under the UN-Women Strategic Plan. The purpose of this framework is to ensure, as far as possible, that the evaluations support UN Women's mission to make progress towards gender equality and women's empowerment. 'The 2022–2025 CEP includes a) 13 corporate evaluations to be led by IES; b) approximately 30 country portfolio and regional evaluations that are strategic independent assessments led by IES; and c) technical assistance to decentralized evaluations and headquarters-led evaluations conducted by independent external evaluators and managed by programmatic offices.'²⁶ The evaluation handbook provides clear guidance on how to approach these diverse types of evaluation.²⁷

Corporate evaluations

Corporate evaluations are 'selected based on the parameters and requirements set by UN-Women's Evaluation Policy and Evaluation Coverage Norms, and through a consultative process with UN-Women stakeholders including the Executive Board Bureau, senior management, and IEAS staff.'²⁸ Their purpose is to contribute to improving organizational performance and strengthening the coherence between normative and operational work.

In the past two years, the following corporate evaluations have been conducted by the Independent Evaluation Service:²⁹

2025

- [Independent corporate evaluation of UN Women's approach to violence against women prevention and response: Navigating changes in global and regional contexts](#)
- [Formative evaluation of UN Women's work in sustainable finance](#)
- [Corporate thematic evaluation of UN Women's support to women's participation in peace processes](#)

2024

- [Feminist collaborative evaluation of UN Women's approach to social norms change](#)
- [Corporate thematic evaluation of UN Women's support to women's political participation: Insights from the field](#)
- [Are we getting there? A synthesis of UN system evaluations of SDG 5](#)
- [Meta-synthesis of UN Women evaluations on advocacy and communications](#)

²⁶ 2022–2025 Corporate Evaluation Plan (CEP)

²⁷ UN Women Evaluation Handbook: How to manage gender-responsive evaluation (2022) <https://www.unwomen.org/en/digital-library/publications/2022/05/un-women-evaluation-handbook-2022>

²⁸ 2022–2025 Corporate Evaluation Plan (CEP)

²⁹ <https://www.unwomen.org/en/about-us/accountability/evaluation/corporate-evaluations>

In terms of coverage, as reported to the Executive Board in June 2025, the coverage of evaluations at UN Women is broadly aligned with the Corporate Evaluation Plan 2022-2025.³⁰ Our interviews identified a range of ways in which they suggested coverage should evolve: ex ante evaluations should be increased for new initiatives, real-time evaluations prioritized for challenging programmes, monitoring focused on routine programmes, and resources freed up for understanding long-term, sustainable transformation in the lives of women and girls. These suggestions are helpful and not out of line with current thinking among decision-makers we interviewed, and merit further consideration of how to ensure uptake.

Despite the consultative process identifying what should be evaluated, some respondents noted that the process by which topics are selected for corporate evaluations may also warrant closer scrutiny. Questions were raised about who has influence over the final list of evaluation topics, whether the choices adequately reflect areas of greatest strategic need or risk, as well as the degree to which the subjects are sufficiently flexible to allow for changes to be made if needed and for evaluations to be timed in a way that they inform key processes and allow real world progress relevant to the implementation of results.

Decentralized evaluations

The large majority of UN Women's evaluations are managed by regional, multi-country, and country offices in partnership with national stakeholders. There were fifty evaluations conducted at the decentralized level in 2024, including 18 strategic evaluations at the country and regional levels led by IES.³¹ The decentralised evaluations are managed by regional, multi-country, and country offices (along with stakeholders), delivered by independent evaluators, and with IES providing technical assistance and quality assurance support. Evaluations include thematic, programmatic, and project evaluations. At the decentralized level IES led regional, thematic and country portfolio evaluations, conducting eighteen country and regional portfolio evaluations and completing six within the year. Concerns were raised by interviewees about the large number of evaluations, limited funding provided for the conduct of each assignment, uneven expertise of evaluators recruited to deliver evaluations, the timing of the assignments vs what can be secured from them, and a wider concern was that there was insufficient attention paid to finding common gaps and key recommendations.

As described above, the quality assurance processes for the decentralized evaluation system are primarily delivered through the Global Evaluation Reports Assessment and Analysis System (GERAAS³²). GERAAS was established by IES in 2013 and it was designed particularly to strengthen quality assurance at the decentralized level. The GERAAS Evaluation Quality Assessment matrix aims to strengthen the quality and credibility of evaluations and are aligned with UNEG norms and standards. It was noted among interviewees that GERAAS might constrain innovative approaches to evaluation, and that by focussing on each individual evaluation cross cutting weaknesses (in relation to identifying common gaps and key recommendations) might be missed.

³⁰ Annual session of 2025 17–19 June 2025 Item 3 of the provisional agenda Evaluation Report of the evaluation function of the United Nations Entity for Gender Equality and the Empowerment of Women, 2024

³¹ Report on the annual session of 2025, 17 to 19 June 2025 Second regular session of 2025 9–10 September 2025 Item 1 of the provisional agenda Organizational matters

³² <https://www.unwomen.org/sites/default/files/2022-06/Evaluation-GERAAS-guidance-2021-en.pdf>

The quality of decentralized evaluations reflects human resources, culture, management and finance. The 2020 OIOS inspection³³ found that decentralized evaluations were generally of a lower quality than corporate evaluations and this was not disputed in our interviews. IES provides additional support in the form of a handbook on managing gender-responsive evaluations,³⁴ e-learning, and guidelines on portfolio evaluations.³⁵ In relation to human resources; people hired to deliver external evaluation were said to be 'variable' by more than one key informant. In the management of decentralized evaluations, it was noted by one interviewee that the 'considerable' management turnover meant that institutional memory could be lost. Concern was expressed that the 'pivot to the field may put more strain on decentralized evaluations in the short term'.

In February 2022 the Executive Director introduced the concept of a pivot to the field as one of her three key priorities emphasising that this would involve strengthening efforts 'to recalibrate and re-capacitate our presence in our country offices'³⁶. This will have implications for how the evaluation policy is managed at the local level. Furthermore, efforts to reinforce the role of UN agencies in supporting and advising (as opposed to doing) might have implications for what is evaluated and how findings and recommendations are communicated. As this looks to the future it goes beyond the remit of this Report but it should continue to be a subject of discussion at the leadership level.

Finding 6. The culture of evaluation is supportive of what is done well, but findings and recommendations focused on key challenges facing UN Women need to be prioritised, and agreed management responses should be followed up.

Formal structures and guidance are important but so too is the culture in and around the evaluation system. Almost every interviewee emphasised that the environment in which evaluative findings 'land' is one of a high pace of work, a changing development context, and the pressure (including self-imposed) to deliver. Some emphasised the rapidity of change and the difficulty of assessing progress and priorities, wondering whether (in the opinion of some interviewees) the sheer number of recommendations made it hard to prioritise these and track implementation. It was also noted that the nature of the evaluation assessment process makes some aspects of the evaluations very prescriptive, including for example the focus of the recommendations, leading to a possible routinization of what is recommended.

Management, it was said by one interviewee, is 'bombarded' with recommendations (but another interviewee questioned this characterisation). A third respondent added that the aggregation of audit, evaluation and external assessment was 'too much... so busy responding we can't do the implementation...recommendations don't have resources...'. However, there were conflicting views, and one observer had not heard about there being too many recommendations and for this person the problem was that resources did not follow recommendations. Given these differences of opinion, it is hard to draw a firm conclusion other

³³ OIOS Inspection of UN Women Evaluation Function Inspection and Evaluation Division 28 May 2020 Assignment No: IED-19-012

³⁴ <https://www.unwomen.org/en/digital-library/publications/2022/05/un-women-evaluation-handbook-2022>

³⁵ <https://www.unwomen.org/en/digital-library/publications/2022/10/un-women-country-portfolio-evaluations-revised-guidelines>

³⁶ <https://www.unwomen.org/en/about-us/business-model/strengthening-un-womens-global-presence>

than, for at least some informants, there is a sense that there are shortcomings in how the production of recommendations is hard-wired to informing better management decisions.

Concern over the number of recommendations may wrongly focus on the quantity rather than the culture around how management and implementing teams use the findings and recommendations. It was suggested to us that having internal sessions to strengthen learning after the evaluation is completed should be (but is not always) routine, and when they do take place there are often insufficient resources to ensure uptake. This highlights the need to continue to promote a culture where evaluation is understood as an opportunity to grow, develop and improve. This calls for more attention to be paid to management culture, on the one hand, and much more focus on dissemination of actionable evaluation findings, on the other. More widely, the value of evaluation comes both from not only immediately acting on recommendations, but also from the secondary impact on how programme staff and management use it to critically examine their assumptions and behaviours. In addition to the questions raised about the number of, and response to, recommendations, it was widely reported that very few people would read a report of 50 pages or more. This is an issue recognised by the IES leadership.

The challenge for the implementation of the evaluation policy at UN Women is that it must both ensure high professional evaluation standards, and be responsive to institutional priorities and the delivery context which it seeks to improve. A further, potentially growing, challenge for the future is that UN Women is increasingly expected to work jointly with other agencies, with national governments, NGOs and the private sector, while at the same time demonstrating measurable impacts on the lives of beneficiaries. Interviewees had several ideas about how to respond to these challenges.

We fully recognise that decentralized evaluations are often donor mandated, and this limits how decentralized evaluations are conducted. However, within this constraint, suggestions were made. All of these would require engagement with donors and encouragement for them to support such changes. One such response was that the current approach was too focussed on evaluating programs and too concerned with describing what happened as opposed to understanding root causes that may be barriers to achieving the corporate strategy. It was suggested that finding the common gaps and key recommendations across evaluations would increase the utility of evaluations. It was also suggested that there are two ways to strengthen decentralized evaluations. The first would be to focus formative evaluation (ex-ante and midterm) on new approaches, not previously tried, rather than evaluating relatively well understood interventions or ones that will be discontinued because resources do not allow for their continuation irrespective of evaluation findings. The second is to have fewer, deeper impact evaluations 'to demonstrate concrete reach on the ground'. Such formative evaluations, we were told, might also help avoid the risk of poorly evidenced programmes which are not evaluable. Thus, there would be an evaluative culture (i.e. using evaluative thinking to support evidence informed learning and decisions) as much as a culture of delivering evaluation studies. This would be especially relevant to the pivot to the countries and the regions mentioned above. Giving confidence that this shift could succeed, is the widely recognised belief that the evaluation function was 'invited to the [management] table' and its views were actively sought.

Finding 7. Independence of the evaluation function is respected and supported, especially at corporate level

The formal governance arrangements outlined above emphasise the independence of the evaluation function. All interviewees who commented on the independence of the evaluation function agreed that IES played an active and effective role in protecting it within UN Women. This was seen to be critical to the credibility of the evaluation function. However, one informant suggested this was more 'tangible' at the corporate level in New York but less so at the decentralized level where the walls between program and evaluation functions can be more porous and where there was more direct pressure from donors and programmatic colleagues. Corporate evaluations were seen to be responsive to management needs, and findings were viewed to be relevant and balanced without being (it was commented) 'finger-wagging' or unduly aggressive.

Successfully implementing the evaluation policy requires navigating the relationship between independence and utility. IES, as custodians of evaluation policy, has navigated this relationship with skill, but the changing context will demand even more creative thinking, leadership, and engagement from senior management.

Finding 8. The implementation of the evaluation policy does not sufficiently support how UN Women provides a narrative of their impact

Evaluation is changing. In the wider profession, there is less focus on 'did the programme work' and more on 'what does this tell us about how to transform delivery and people's lives?'.³⁷ Beyond UN Women, evaluation as a set of practices is undergoing significant changes. As UN Women undertakes its pivot to the countries and regions, it will be greatly strengthened by drawing on the rise of evaluative thinking – that is, combining real time evaluation with innovation, learning, adaptation, citizen-responsiveness, and agile management. According to OECD, 'Evaluative thinking contributes to new learning by providing evidence to chronicle, map and monitor the progress, successes, failures and roadblocks in the innovation as it unfolds. It involves thinking about what evidence will be useful during the course of the innovation activities, establishing the range of objectives and targets that make sense to determine their progress, and building knowledge and developing practical uses for the new information, throughout the trajectory of the innovation.'³⁸ The rise of digital technologies and AI will both reinforce this and create new opportunities and challenges.

MOPAN (p.15) notes: 'one of the most significant challenges UN Women faces is ensuring that its interventions generate long-term, sustainable change rather than isolated, project-based successes.'³⁹ Evaluation practice at the corporate level already partly reflects this and must increasingly support insights around how best to deliver long-term, sustainable change. This is a challenge also for Results Based Management and monitoring.

It is important to see the implementation of the evaluation policy as a key part of delivering long term change for women and girls and not a detached function. Amid growing uncertainty, driven

³⁷ The Future of Evaluation is discussed in a World Bank IEG YouTube from April 2024 at <https://www.youtube.com/watch?v=CyglGIFUOs&t=71s>

³⁸ Lorna Earl, Helen Timperley **Evaluative thinking for successful educational innovation** OECD <https://doi.org/10.1787/5jrxk1jtdwf-en>

³⁹ MOPAN (2025), MOPAN Assessment Report: UN Women (Part I), MOPAN, Paris © MOPAN 2025.

by funding challenges, global fiscal constraints, political dynamics, and ongoing transformation, it is important to align multi-year evaluation plans annually to ensure alignment with the highest priorities and stakeholder needs. The Advisory Committee on Oversight commends IES's agility in this regard, demonstrated through its rapid formative assessment of UN Women's Gender Equality Accelerators (GEA), providing senior management with timely and relevant information regarding ongoing transformation.⁴⁰

The need to embed evaluative thinking into strategic planning is clear. Evaluative thinking involves not only using evaluation findings to improve discrete programme-related choices but a wider willingness to use data and outcomes to reflect on the skills, behaviours and knowledge that drive better outcomes. KPIs support the tracking of performance over multiple years and provide a focus for management to ensure accountability. They are fundamental to supporting the operation of the evaluation function, but should be viewed as a springboard for further understanding rather than definitive proof of success. For this, the evaluation system must be able to both tell and inform a strong performance story⁴¹ using numbers to connect programme achievements to the corporate objectives of UN Women. The ACO, while highlighting 'the importance of strengthening UN-Women's evaluation capacity and ensuring the effective use of evaluation findings and recommendations at the regional and country levels' (p. 5) also noted the need to strengthen further its ability to convey UN Women's 'performance story'.

Measuring progress of UN Women in general and the impact of their activities in particular provoked considerable discussion among interviewees. It was recognised that gender equality change can be very hard to measure and some felt that the demand for more quantified impacts was not always based on an understanding of how challenging this is. On the other hand, it was stressed that in the current environment, demonstrating results will be key to securing funding. Furthermore, programs which are implemented without a clear route to impact, measurable indicators, and routine monitoring data will prove hard to evaluate and will not yield useful lessons for management and other stakeholders and, as one interviewee suggested, 'it kills donor enthusiasm'. Expectations for measuring impact should be cautious and modest but nevertheless should be done – often using a Contribution Analysis. Engaging with, and incorporating, evaluative thinking may be among the greatest challenges facing the implementation of evaluation policy in UN Women. Contribution Analysis⁴² allows a structured approach to both tell a performance story and identify where innovations and improvements might be needed.

Finding 9. High-scoring KPIs for the use of evaluation and management response only tell half the story

Interviewees were especially forthcoming about ways to strengthen the use of evaluations, suggesting there was a continuing sense that improvement was needed despite relatively high percentages given to management acting on recommendations (Table 1). In order to improve further the utilisation of evaluations, several areas should be considered.

⁴⁰ Report of the Advisory Committee on Oversight for the period from 1 January to 31 December 2024
[file:///C:/Users/tling/Downloads/UNW_2025_3_Add.1-EN%20\(1\).pdf](file:///C:/Users/tling/Downloads/UNW_2025_3_Add.1-EN%20(1).pdf)

⁴¹ <https://www.betterevaluation.org/tools-resources/reporting-outcomes-setting-performance-expectations-telling-performance-stories#:~:text=Key%20features,done%20to%20assure%20quality%20data?>

⁴² [Contribution analysis | Better Evaluation](#)

1. **Timeliness:** As is often repeated in the wider literature,⁴³ the timeliness of evaluations is as critical to their usefulness and efforts have been made to cut the time taken on evaluations while at the same time seeking to reduce the length of evaluation reports with a widespread view (shared by the leadership of IES) that reports over 50 pages long would not get read, and that senior management preferred even shorter reports. The key issue is to understand how the timing of evaluations fits with decision making cycles and needs.
2. **Deliverables/Communication:** Reinforcing efforts to communicate deliverables in a variety of ways (for example, more use of magazine formats such as *Transform*⁴⁴) would strengthen utilisation. One interviewee felt that findings were not presented in the way that makes a difference to management, with insufficient critical edges and tending to be too 'nice'). Systematically, blunt messages (for example what should be stopped) are much less common in UN Women than positive affirmation.
3. **Recommendations and learning:** It was suggested that fewer, more graduated and prioritised recommendations would gain more traction. The wider literature, including from UNEG, is equally clear.⁴⁵ Recommendations should: prioritize for impact; be manageable; be specific; and based on the evidence presented in the evaluation report. However, as discussed above, we suggest that what is equally important is the culture of learning and support that lead to recommendations having positive benefits for stakeholders. This is as much the responsibility of management as of IES. In relation to such learning and support, there also a contrast to be drawn between the MOPAN scores for KPI 8 'Evidence-based planning and programming' which scores 'independent evaluation function', 'evaluation coverage', and 'evaluation quality' all very highly but scores 'poor performance tracked' as 'unsatisfactory'. For wider organisational learning, it may be that the evaluations meet the criteria for quality but that poor performance tracking – which would allow failure to adopt agreed recommendations to be censured – prevents consistent learning.
4. **Focus/Theme:** Evaluations focus too much on what has happened rather than how it happened. Interviewees wanted more on how to design, manage and deliver change rather than describing what a programme did. Similarly, it was felt that there were too many programme or project evaluations which do not explore the root causes standing in the way of UN Women delivering change for women and girls. At the strategic level, meta-findings that create insights by combine and consolidate evaluation findings are seen to be 'most useful' especially where these highlight common trends. The short paper on Gender Equality Accelerators was highlighted as an example of a shorter and different kind of product that could be impactful.

⁴³ <https://researchonline.lshtm.ac.uk/id/eprint/4652055/3/Timely-evaluation-in-international-development.pdf>

⁴⁴ For example, [TRANSFORM – The magazine for gender-responsive evaluation – Issue 31, September 2025 | Publications | UN Women Knowledge portal](#)

⁴⁵ See, for example, United Nations Evaluation Group. (2018). Improved quality of evaluation recommendations checklist. Retrieved from <http://www.unevaluation.org/document/detail/2124> This checklist includes useful background information, and it also deals with the follow-up of evaluation recommendations, using the UNEG Good Practice Guidelines for Follow-up to Evaluations (<http://www.unevaluation.org/document/detail/1412>). See also: Feinstein O 'Checklist for Evaluation Recommendations' <https://files.wmich.edu/s3fs-public/attachments/u3036/2019/eval-rec-feinstein.pdf#:~:text=Number:%20Limit%20the%20number%20of%20recommendations.%20Be,as%20this%20jeopardizes%20follow%20up%20of%20their%20implementation.>

Meanwhile at the decentralized level, and often a consequence of donor requirements, interviewees highlighted an over-emphasis on ex post evaluations when the programme had not been designed with a clear theory of change and measurable outcomes, and the management information and monitoring did not support an evaluation. This is a fundamental problem of evaluability.⁴⁶ Unevaluable programmes are unlikely to support insights for future improvement. This, of course, is a problem of program design rather than the responsibility of evaluators in general or the IES in particular. Also, as discussed, at the decentralized level, evaluations must provide donors with the information they demand in return for their investments. These requirements may or may not align with the characteristics of delivering high impact evaluations and may focus on projects which will be discontinued and where lessons cannot be capitalised upon). Thus, at the decentralized level, it was said there is a 'whole conveyer belt of things that need to happen'. Improving this conveyor belt is a shared responsibility across management, donors, decentralized evaluators and IES.

5. **Selecting high impact evaluations:** There was little disagreement about what makes for high impact evaluations and the list of characteristics included: management interest; timing; quality of the evaluation; having clear impacts; presentation; and novel insights). However, we found little evidence that these criteria determine the choice of evaluations or the priority given to one topic over another. At the same time, we note that high impact evaluations may require longer studies and reports; a requisite variety of evaluation types, lengths, and foci will continue to be required and this should reflect the strategic aims of evaluation as discussed in this Report.

Overall, there was broad support for the idea of moving away from a heavy reliance on more traditional approaches to conducting and reporting evaluation, and providing more reviews, short reports and so forth. However, there was no suggestion that 'traditional evaluation' (in terms of length, focus, and timeliness) had no place in the future; this requires evolution that builds on existing strengths and not revolution.

Finding 10. UN Women makes a significant, if hard to measure, contribution to the UN-wide evaluation systems and national gender-responsive evaluation capacity

IES promotes gender responsive evaluations across the UN system through its leadership in UNEG working groups, co-chairing support for the EvalGender+ initiative, contributing to UN system-wide frameworks, and advocacy for gender responsive evaluation in the UN system. IES contributed to the SDG Synthesis Coalition to synthesize evidence on SDG achievements and led the first joint synthesis of UN evaluative evidence on SDG 5. IES also supported several United Nations Sustainable Development Cooperation Framework (UNSDCF) evaluations, and contributed to the Global Evaluation Initiative's efforts to strengthen national evaluation capacity. In 2022, IES joined the Inter-Agency Humanitarian Evaluation group which assesses the results of the Inter-Agency Standing Committee members' collective humanitarian response and actively supports a strong gender perspective in evaluation.

⁴⁶ https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf

UN Women, and IES in particular, plays a positive role in supporting gender responsive evaluations across the UN system as well as supporting rights-based approaches and systems thinking, although the impact of this is hard to measure. Interviewees considered this to be an effective way in which UN Women could use its expertise and influence to support positive change, and contributing to the UN-wide system might merit being allocated even more resources. It was also speculated that given its importance the UN-wide role should receive much more senior management time and the system wide evaluation function needed strengthening and made more coherent.

UN Women, and IES in particular, also support national gender-responsive evaluation capacity development through technical assistance, training, and partnerships with national and regional organizations to help governments and other bodies conduct evaluations that promote gender equality.⁴⁷ 'Capacity development has been a long-term intervention strategy within the UN system. It is also a key intervention strategy in UN Women, providing support to duty bearers and rights holders in both technical and functional capacities.'⁴⁸ This support is a key part of the organization's strategy to strengthen national evaluation systems and will become increasingly important as UN Women focus more on influencing others and pivoting to the field.

There are significant emerging challenges to successful implementation of the evaluation policy

The previous section describes a well-structured evaluation system, with good quality guidance and support for managers and evaluators, a quality assurance and reporting system which is systematic and transparent, a willingness to learn from external reviews, positive engagement with other UN entities, and a determination to contribute to the wider learning, policies and strategic orientation of UN Women. All of this supports our conclusion that the implementation of the evaluation strategy at UN Women is maturing and is to a considerable extent effective. We have highlighted above findings related to how policy is currently being implements. In this section we focus on likely future challenges to successful implementation.

Emerging challenges will most likely reshape the operating environment in which UN Women's evaluation system must function. These include the strategic shift toward decentralisation and regional hubs, such as the shift to Nairobi and Bonn, which has implications for resource allocation, capacity, and oversight. As more personnel are based where operational costs are lower and virtual collaboration increases, evaluation systems must adapt to ensure quality and coherence across dispersed teams.

Financial pressure is intensifying. Evaluation must now operate in an environment where core funding is constrained, demand for results is rising, and there is an increasing emphasis on doing more with less. As one senior interviewee noted, *"The ability to tell powerful results stories is life or death [for UN Women], and evaluation is critical for this."* However, this also depends on how programmes are designed and implemented, and not only on the approach taken by IES.

⁴⁷ Executive Board of the United Nations Entity for Gender Equality and the Empowerment of Women 2020 Evaluation policy of the United Nations Entity for Gender Equality and the Empowerment of Women Annual session 2020 22–23 June 2020 Item 6 of the provisional agenda

⁴⁸ <https://www.unwomen.org/en/digital-library/publications/2023/02/evaluation-capacity-development-of-partners-to-respond-to-the-needs-of-women-and-girls-at-national-level#:~:text=Capacity%20development%20has%20been%20a,decision%2Dmaking%20and%20organizational%20learning.>

There are also broader system-level shifts: a growing expectation for collaboration and coherence across UN agencies (already under way and evaluated⁴⁹), a diminishing appetite for a fragmented UN, more collaboration with governments and private sector, more efforts to influence rather than 'do', and the rise of cross-cutting initiatives that require "joined-up" evaluation approaches. The demand for real-time learning and faster feedback loops is growing. This environment will call for a variety of types of evaluation, most likely requiring more that are agile, targeted, and responsive. These would produce reports that are, short, sharp pieces that feed directly into adaptive decision-making.

In addition, current global dynamics include the rapid development of AI, evolving models of development cooperation, growing questioning of multilateralism in some quarters, and the growing recognition of limited progress on the SDGs. Addressing these will also influence how we deliver evaluations and organise evaluation systems. Within this shifting terrain, a key question will concern UN Women's comparative advantage, and how evaluation can help clarify and strengthen this. Evaluation, therefore, must evolve as both a learning and an influencing function.

All of this points to the need for a refreshed evaluation model, one that matches the pace and complexity of UN Women's evolving business model. Fortunately, many of the foundations are in place: strong governance, trusted relationships, solid technical expertise at corporate level, and continued support from the Board for maintaining a 2–3% investment in evaluation. However, engaging with new challenges will require recalibrating the balance of the evaluation portfolio, reframing the focus of evaluations, and aligning with the realities of a more dynamic, collaborative, and resource-constrained environment.

Conclusions: 6 Opportunities for Progress

As is clear throughout this report, the implementation of the evaluation policy at UN Women is maturing and there are very strong foundations to build upon as implementation responds to new opportunities and a changing environment. Therefore, we are proposing a rebalancing, refreshing, and re-orientation which comprises nine evolutions, each doing less of some things and more of another.

For decentralized evaluations:

1. From ensuring routinised coverage or programme description and immediate effect to focussing on evaluative evidence that focuses on root causes and or can help transform the lives of women and girls, and identifies the root causes inhibiting the further progress of UN Women, along with potential solutions. **Responsible agent:** IES must further engage with regional offices to prioritise evaluations and in doing so revisit the balance in the portfolio of assignments, and increase the impact of individual assignments.
2. From multiple routine recommendations to focused and prioritised areas for improving organisational performance that are socialised with management before later being rigorously followed up. **Responsible agent:** regional offices with support from IES as part of the pivot to the regions and countries. Those designing interventions should

⁴⁹ UN Women et al 2024 Getting There Together A Synthesis of Inter-Agency Programme Evaluations on SDG 5

ensure the adequate development of the ToC and the evaluation manager must be very clear on what areas of learning are expected. This would include having a clear understanding of what is already known and feasible and areas that require more learning and recommendations that may allow UN Women to progress.

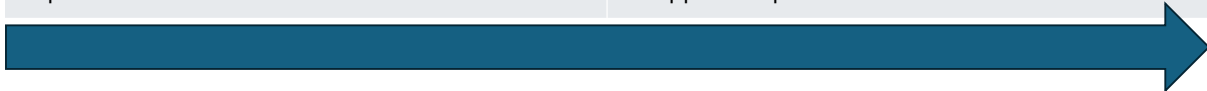
3. From individual donor-led requirements to collaborative agreements with donors and rights-holders about what would most usefully be evaluated. **Responsible agent:** Project Managers and donor engagement staff should discuss the assessment options with donors before projects start to ensure that expectations are managed and requirements met.
4. From disparate knowledge functions to integrated insights (monitoring, audit, evaluation, management information etc), possibly facilitated by digital and AI, that tell and quantify the performance story of UN Women. Often, the use of experimental and quasi-experimental methods will not be possible (for either ethical or technical reasons) and in this situation using approaches that adhere to the principles of Contribution Analysis should be considered as a means to tell an evidence-based contribution story across UN Women's evaluations. **Responsible agent:** IES and senior managers should emphasise the need for clear and evidence-based contribution stories and this should be taken forward by Regional Directors and promoted as part of the pivot to the countries and regions.

At corporate level:

5. From focusing on what is done well to highlighting what needs to change. **Responsible agent:** UN Women Senior management, Programmatic departments must identify issues.
6. From ensuring the independence and support for the evaluative function to more conscious alignment with learning, agility and informing strategy at the corporate level while mobilising change at the decentralized level to enable a successful pivot to the regions and countries. **Responsible agent:** UN Women Senior management, regional leadership.

Table 2 Summary of opportunities for progress.

From.....	To.....
Large volumes of widely spread evaluative evidence describing what happened	Streams of inter-connected evidence identifying root causes and focussed on key solutions to achieving progress for women and girls
Recommendations unevenly followed up	Recommendations focused on key actions that are socialised fully with management and systematically followed up
Responding to donor requirements	Helping donors to see the merits of wider evaluative learning while providing them with a clear impact story
Supportive tone to encourage engagement	More direct tone to unambiguously identify what needs to change
Robust methods used to tell the ‘delivery story’	Where experimental methods are not feasible, focus on telling an evidence-based ‘contribution story’
Strong record of independence and professionalism at corporate level	Ensuring high professional standards across all levels to support the pivot to the field



Annex 1: How the 10 findings relate to the key questions in the ToR on relevance, independence and coverage

Finding Number	Relevance	Independence	Coverage
1	Steady progress towards KPIs reflects alignment with professional standards and stakeholder needs, but the true meaning and impact of KPIs may need further scrutiny.	–	Coverage and financial investment slightly below targets; management response implementation rate lowered after tracking system was updated.
2	Governance arrangements support strategic relevance but require more senior engagement for data quality, learning, and integrating evaluative thinking with programs.	Reporting lines support some independence, but engagement of Executive Board (EB) is critical for high-level decisions and accountability.	Persistent gaps remain in data quality, monitoring, and evaluation coverage; field-level coverage and lesson integration need strengthening.
3	Professional standards grounded in UNEG—clear benchmarking and relevance to gender-responsive priorities.	UNEG engagement strengthens robust independence and professional quality at corporate level.	UNEG frameworks bolster coverage norms, but concerns about decentralized level quality and HR expertise persist.
4	External reviews validate relevance; IES is responsive to recommendations and shows learning culture.	External scrutiny (MOPAN, ACO,) supports independence, with management positively engaging with external feedback.	KPI tracking and budget targets met, but continuing need to systematize learning, validate coverage, and ensure recurring issues are tackled.
5	Systematic evaluation processes (corporate & decentralized) support relevant coverage of Strategic Plan and DAC-OECD criteria.	Selection of evaluations follows policy and stakeholder consultation, with IES guidance supporting balanced independence.	Coverage norms applied systematically; quality assurance via GERAAS and GATE, but decentralized evaluations face challenges: uneven expertise, limited funding, recurring gaps.
6	Culture supports learning and responsiveness, but findings/recommendations need sharper focus for greatest relevance and utility.	Evaluation function is invited into management, supporting independence, but risk of recommendations overwhelming resources or not being used.	Need for stronger follow-up and prioritization for coverage—internal sessions and overarching gap analysis highlighted.

Finding Number	Relevance	Independence	Coverage
7	Interviewees regard independence of function as key for credibility and relevance, especially at corporate level.	Independence respected at corporate level, but more porous/compromised at decentralized level due to donor/program pressures.	Decentralized evaluations less protected, risk coverage gaps with porous boundaries; independence seen as driving coverage quality.
8	Policy insufficiently supports UN Women’s impact narrative and ability to demonstrate measurable results.	–	Calls for proper impact evaluation; RBM and monitoring gaps at the field level diminish coverage/impact measurement.
9	KPIs and management response rates do not fully describe evaluation use or learning; timeliness and format impact relevance for decision-makers.	Open culture and improvement commitment positive for independence but may limit accountability/rigour.	Coordination and focus in decentralized evaluations critical for stronger coverage; implementation remains variable.
10	UN Women, specifically IES, makes a hard-to-measure but important contribution to system-wide gender-responsive evaluation—relevant for UN progress.	–	Amplifies UN system coverage, though impact hard to measure due to challenges in coherence, collaboration, resource, and evaluation function strength.