PERFORMANCE INDICATOR
04

EVALUATION

04. Performance Indicator: Evaluation		
Approaches requirements	**Meets requirements**	**Exceeds requirements**
4a. Meets some of the UNEG gender-related norms and applies some of the standards in the 2024 UNEG Guidance on Integrating Human Rights and Gender Equality in Evaluation	4bi. Meets the UNEG gender equality - related norms and standards **and** 4bii. Applies the 2024 UNEG Guidance on Integrating Human Rights and Gender Equality in evaluation during all phases of the evaluation	4ci. Meets the UNEG gender equality - related norms and standards **and** 4cii. Applies the 2024 UNEG Guidance on Integrating Human Rights and Gender Equality in Evaluation during all phases of the evaluation. **and** 4ciii. Conducts at least one evaluation to assess the entity's corporate performance on gender mainstreaming or equivalent every 5years

What is the Evaluation indicator?

The UN-SWAP Evaluation Performance Indicator assesses the extent to which the evaluation reports of an entity meet the gender-related UNEG Norms and apply the Standards and demonstrate effective use of the 2024 UNEG Guidance on integrating human rights and gender equality during all phases of the evaluation. It also calls on all reporting UN system entities to conduct at least one evaluation to assess the entity's corporate performance on gender mainstreaming every 5 years. This might constitute, but not be limited to, corporate evaluation of gender policy, mainstreaming, and strategy or equivalent".

A well-functioning evaluation system is intrinsic to good project/programme management, in terms of ensuring both accountability and success for the achievement of the 2030 agenda for sustainable development by 2030.

In addition, evaluation is a relevant tool for learning and accountability. Evaluations generate lessons learned and knowledge for future decisions and timely corrections on projects/programme implementation.

UNEG gender-related Norms, Standards and Guidance

The UNEG Norms and Standards for Evaluation were updated in 2016 and for the first time, included a stand-alone Norm on human rights and gender equality. The new Norm on human rights and gender equality calls on evaluators and evaluation managers to ensure that these values are respected, addressed and promoted, underpinning the commitment to the principle of 'No-one left behind'. Taking into consideration that the UNEG Norms and Standards are the normative framework that guided evaluation policies and guidance of UN entities, the adoption of the stand-alone Norm on Human Rights and Gender Equality is paramount for ensuring it is institutionalized across the UN system.

![Gears icon] **How to use this performance indicator**

The use of the UNEG endorsed [UN-SWAP Evaluation Scorecard](link) provides a basis for harmonization across entities by assigning an overall aggregate score for reporting against the UN-SWAP Evaluation Performance Indicator's scaled rating system: missing, approaching requirements, meeting requirements, or exceeding requirements.

An entity should only report 'not applicable' if there is no evaluation unit or evaluations conducted by the entity. However, if no evaluations were conducted in the previous year, the last rating completed should be used with a clear note indicating the year upon which the rating is based. This approach is being used to avoid confusion with those entities that do not have an evaluation unit/conduct evaluation and thus the indicator is 'not applicable'.

Essential steps of the UN-SWAP EPI reporting cycle
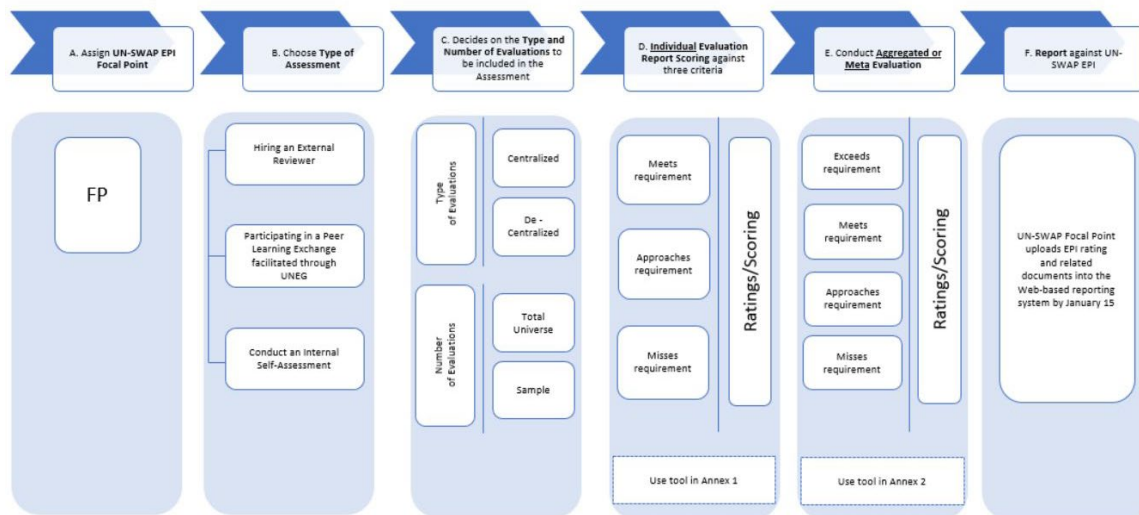
What should be included in the UN-SWAP meta-review/evaluation?
For the purpose of reporting against this indicator, UN entities should include in their UN-SWAP meta-review/evaluation only those reports that meet the UNEG definition for evaluation. Although there are some exceptions, the evaluations included should have been finalized in the period being reported: annual reporting cycle January – December.

How many evaluation reports should be assessed for the UN-SWAP EPI reporting?
Evaluations conducted or managed by central evaluation offices and decentralized evaluations can be included. The general recommendation is to include all centralized/corporate evaluation reports. Entities with a decentralized evaluation function can either include the total universe of decentralized evaluations or a sample of evaluations, accurately reflecting the different types of evaluations. Those entities with established quality assessment and meta-evaluation systems are encouraged to include the total universe of evaluations for the year under review.

Those entities selecting a sample of evaluations for meta-review/evaluation should aim to select a representative sample so as to minimize sample bias.

The UN-SWAP Evaluation Scorecard

The UNEG endorsed scorecard is a tool aimed at assessing evaluation reports of an entity against three criteria. Through its fourth criterion, the scorecard also calls on all reporting UN system entities to conduct at least one evaluation to assess corporate performance on gender mainstreaming every 5-8 years[10].

The first two criteria look at whether gender equality concerns were integrated in the evaluation scope of analysis and methods and tools for data collection and analysis.

- GEEW is integrated in the evaluation scope of analysis and evaluation criteria and questions are designed in a way that ensures GEEW related data will be collected.
- A gender-responsive methodology, methods and tools, and data analysis techniques are selected.

The third criterion is focused on whether the evaluation report reflects a gender analysis as captured in the findings, conclusions and recommendations – this could be captured in various ways throughout the evaluation report.

- The evaluation findings, conclusions and recommendations reflect a gender analysis.

The fourth criterion is focused on whether the entity has commissioned:

- At least one evaluation to assess corporate performance on gender mainstreaming or equivalent every 5 years.

The scope and title of evaluations to assess corporate performance on gender mainstreaming differs from entity to entity. This might constitute, but not be limited to, corporate evaluation of gender policy, mainstreaming, and strategy or equivalent".

How to score each evaluation criteria?
UN entities will use the UNEG endorsed UN-SWAP EPI Scorecard to assess each evaluation report using a four-point scale (0-3) rating system for each criterion (tools are provided in the UNEG endorsed Technical Note and Scorecard). Each of the scoring levels below corresponds to a numbered score:

0 = Not at all integrated. Applies when none of the elements under a criterion are met.

1 = Partially integrated. Applies when some minimal elements are met but further progress is needed and remedial action to meet the standard is required.

2 = Satisfactorily integrated. Applies when a satisfactory level has been reached and many of the elements are met but still improvement could be done.

3 = Fully integrated. Applies when all of the elements under a criterion are met, used and fully integrated in the evaluation and no remedial action is required.

It is important to note that decimals should not be used in the scoring of criteria; use only whole numbers.

How to score individual evaluation reports?
After reviewing the individual evaluation report for each criterion, a score is assigned as follows:

---

[10] The scope and title of evaluations to assess corporate performance on gender mainstreaming or an evaluation of gender equality policy/strategy differs from entity to entity. This might constitute but not limited to corporate evaluation of gender policy, gender mainstreaming strategy, plan or equivalent.

- 0-3 points = Misses requirement
- 4-6 points = Approaches requirement
- 7-9 points= Meets requirement

Since <u>each evaluation report</u> is assessed against three criteria, the maximum number of points that a report can obtain is 9 (3 points for each criteria). For example, if the score is 7 or above, the rating for the evaluation report would be "Meets Requirement". Use Tool provided in Annex 1 of the <u>UN-SWAP EPI Technical Note.</u>

How to calculate the meta-score?

Once you have filled in the scorecard for each individual evaluation (which requires a new worksheet in the excel spreadsheet) you are ready to calculate the aggregate score in the meta-evaluation scoring sheet. Scores for each individual evaluation are added up and divided by the total number of evaluation reports reviewed (see Annexes 2 and 3).

- 0-3,49 points = Misses requirement

- 3,50-6,49 points = Approaches requirement

- 6,50-9,0 points= Meets requirement

- 9,01-12 = Exceeds Requirement

Please note that when using the UNEG UN-SWAP Scorecard and the Meta-Score Calculation Tool, the additional 3 points for conducting a corporate gender evaluation are to be added only if the average score for evaluations is 6.5 or higher, i.e. in cases where an entity is "meeting requirements" based on the average evaluation score. The additional 3 points cannot, for example, be added to an average score for evaluations that falls between 3.5-6.49 or "approaches requirements" to bring the score up to "meets requirements."

As mentioned in Annex 3 of the EPI Technical Note, **the requirement of conducting a corporate evaluation on gender mainstreaming is to be considered under the category "exceeds" only.**

A substantive revision of the UN-SWAP technical note and accompanying scorecard will be undertaken in 2025. The revision will include alignment with relevant frameworks and guidance documents, including the UN-SWAP 3.0 framework, the <u>2024 UNEG guidance on Human Rights & Gender Equality</u>, the <u>2020 UNEG Ethical Guidelines for Evaluation</u>, and other key documents.

Evidence base

Examples of documents to attach to substantiate the entity self-assessment for this indicator:

- Aggregated/meta-evaluations
- Completed UN-SWAP Evaluation Scorecards
- Report of corporate gender mainstreaming evaluation
- Management response to the corporate gender mainstreaming evaluation

Note: Please identify a self-explanatory title for the documents uploaded onto the platform, particularly for those shared to the UN-SWAP Knowledge Hub.

Who is the Business Owner for this indicator?

The adequate establishment and reporting on gender-responsive evaluations would be the primary responsibility of an entity's evaluation unit. Generally speaking, the entity's evaluation unit is expected to be a member of the United Nations Evaluation Group (UNEG) and adhere to the UN norms and standards for evaluation.

Elements derived from the Gender Equality Acceleration Plan (GEAP)

✓ The Secretary General will request OIOS to include gender equality indicators emerging from this Gender Equality Acceleration Plan in its standard operating procedures to assess entities' capacity, leadership performance and impact when conducting independent assessments UN of entity/ departments.

✓ The Secretary General will further request the UN Evaluation Group to review its evaluation standards in light of this Acceleration Plan.

✓ All other entity heads will engage their evaluation and audit departments to ensure that measuring progress against entity gender equality acceleration plans are reviewed and reported on to their governing bodies.

What are the specific requirements for this indicator?

How to approach requirements

To approach the requirement for this Performance Indicator, at least one of the three criteria needs to be assessed at "Satisfactorily Integrated (2)".

How to meet requirements

For an evaluation to "meet requirements" at least one of the criteria needs to be assessed at "fully integrated (3)". For example, if there are three evaluations in the meta-evaluation that have individual scores of 9, 8, and 6 respectively, the sum of the three scores would be 23, which divided by 3 (the number of evaluations under review) would give a mean score of 7.6 points. This would give an aggregate rating of *"Meets Requirement".*

At a minimum, each UN system entity should aim to "meets requirement" related to this Performance Indicator in terms of integrating gender equality and empowerment of women (GEEW) in their respective evaluations. However, achieving this is only considered a starting point to fully integrating gender dimensions in evaluation processes, rather than an end in and of itself. UN entities should continually strive to *"exceeds requirement"* if the UN system is to truly benefit from gender-responsive evaluation practice.

Examples of completed Scorecards on how to report against this performance indicator can be found UNEG endorsed *UN SWAP EPI Technical Note* - Annex 1 and 2.

How to exceed requirements

In order to exceed requirements, an entity's evaluation reports must 'meet requirements' and that entity must also conduct an evaluation of its corporate gender policies. In other words, for an entity to "exceed requirements," the aggregate score of its evaluation reports must "meet requirements" by achieving a score of 6,5 points or higher AND it must have conducted a corporate evaluation of its corporate performance on gender mainstreaming. Otherwise, even though an entity conducts a corporate evaluation, but its reports don't meet requirements, its overall score cannot be in the exceed category. Or else, the maximum score for review of evaluation report/s would remain 9, achieving a rating of "meets requirement".

The Guidance on Evaluating Institutional Gender Mainstreaming is a practical guide designed for UNEG members, policy and programme evaluators and others who advocate for a common approach to assessing progress of institutional gender mainstreaming (GM) in the United Nations system. This guide responds to recommendations from the UN Women "Review of Corporate Gender Equality Evaluations in the United Nations System", which was drafted in consultation with UNEG, EvalPartners, and EvalGender, and the revised Evaluation Performance Indicator for the United Nations System-wide Action Plan (UN-SWAP), which identifies a gender equality evaluation as necessary for "exceeding requirements." The aim of the guidance is to encourage UN entities to conduct more evaluations of their respective institution's GM approach to improve results, and to enhance coherence of information across the UN system on progress made towards gender equality and empowerment of women (GEWE).

Important considerations for the exceed requirement

Please note that any corporate gender mainstreaming/strategy/policy or equivalent evaluation conducted within the five years preceding the period being reported is eligible for consideration. If the corporate evaluation was conducted more than five years prior to the reporting period, then it is ineligible for consideration. This means that an entity must have conducted a corporate evaluation within the preceding five years to achieve "exceeds requirement."

For example, the **World Food Programme (WFP)** Office of Evaluation completed an evaluation of WFP's Gender Policy in 2014. This means that conditional on the aggregate score of its evaluation reports "meeting requirements", WFP can add the additional 3 points for the UN-SWAP 2.0 cycle for eight consecutive years, i.e. until the 2021 annual reporting. An entity that completed a corporate evaluation on gender mainstreaming/policy/strategy in 2010, however, would not be entitled to the additional 3 points for its annual UN-SWAP EPI reporting in 2018 and beyond as the entity is due to undertake a new corporate evaluation.

Cognizant of the resource constraints by smaller entities to commission external evaluation of their respective gender policy/strategy, evaluations by internal parties could be considered as adequate to get the additional three points. This is to give those entities that are committed to improve their gender equality policy/strategy the opportunity to reach "exceed requirement".

The addition of the "exceed category" of evaluations on gender mainstreaming or evaluation of gender equality policy/strategy provides an excellent opportunity for validating institutional progress, particularly given that UN-SWAP annual reports are based on self-assessment.

As a way to demonstrate compliance with the "exceeds requirement" criterion of the UN-SWAP EPI and whether the evaluation is still valid (i.e. less than 5 years old), the evaluation report of corporate performance on gender mainstreaming should be uploaded to the UN-SWAP online Reporting System annually.
Examples of a completed Scorecard on how to report against this performance indicator can be found in UNEG endorsed UN SWAP EPI Technical Note - Annex 2.

Online reporting platform and qualitative feedback
During annual UN-SWAP reporting, Evaluation Offices are responsible for conducting and sharing their aggregated/meta-evaluations and/or completed Scorecards with their organizational UN-SWAP Focal Point[11], who is responsible for uploading these to the web-based reporting system. Evaluation Offices are also encouraged to include examples of evaluations that demonstrate how entities are approaching, meeting or exceeding requirements for this indicator overall or for specific dimensions for upload to the web-based system.

The evaluation report of the corporate gender mainstreaming/strategy/policy or equivalent should be uploaded to the UN-SWAP online reporting platform annually as proof of achieving "exceed requirement."

 Example: Meeting Requirements

The **United Nations Population Fund (UNFPA)** Evaluation Office commissioned an independent quality assessment of 22 evaluations covered in the 2017 UN SWAP EPI reporting cycle: 2 corporate evaluations and 20 decentralized programme-level evaluations (including 19 country programme evaluations and one regional programme level evaluation) were included.

The assessment used the UNEG endorsed Technical Note and Scorecard, which specifies the overall score as well as the score by evaluation (vis-à-vis the four criteria comprising the EPI), providing a detailed explanation of why a particular rating was given. In 2017, on aggregate, UNFPA evaluation reports "met the requirements" of the EPI, with an overall score of 9.45, reflecting a year on year improvement from 2015 to 2017 in the quality of evaluation reports, including the integration of GEEW. Twenty of the 22 evaluation reports "met the requirements", while one "exceeded requirements". One report (a corporate evaluation), "approached requirements".

The evaluation quality assurance and assessment (EQAA) system expands and strengthens assurance and assessment processes. The evaluation quality assessment grid against which all evaluations are assessed – includes a criteria on gender which directly mirrors the language of the EPI. In 2017, the grid (and its use) was further strengthened: a guidance note was developed and the approach to addressing the sub-criteria (under each criterion in the grid) was made explicit, ensuring consistency in and transparency of the assessment.
However, challenges to integration continued to be faced, including in ensuring adequate resources are directed toward integration (which often requires a radical re-think in the methodological approach to evaluation), including ongoing practical training on integrated gender equality in evaluation.

---

[11] All UN entities have designated UN-SWAP Focal Points who consolidate reporting against all UN-SWAP performance indicators and that enter the data in the online reporting system on behalf of their respective organizations. These colleagues are generally staff of the Gender Units/Gender Divisions of the entities.

Example: Exceeding Requirements

The **International Fund for Agricultural Development (IFAD)** Independent Office of Evaluation (IOE) undertook a meta-evaluation of 17 evaluations (completed in 2016-2017), comprising of 9 project performance evaluations (PPEs), 1 impact evaluation (IE), 5 country strategy programme evaluations (CSPEs) and 2 evaluation syntheses against the criteria set out in the UN SWAP EPI scorecard. The individual meta-scores across different types of evaluation vary. 12 out of 17 (70%) evaluations exceed requirements, 4 (24%) meet requirements, 1 (6%) approaches requirements, resulting in a meta-score of 10.7. IOE has thus "exceeded requirements".

The highest average score (2.88 out of 3) was on performance criteria 1 (scope of analysis and indicators designed ensuring that GEEW related data be collected) and criteria 2 (2.82 out of 3) (evaluation criteria and evaluation question specifically address how GEEW has been integrated into the design, planning, implementation of the intervention and the results achieved). The criterion scoring the lowest was criterion 4 (findings, conclusions and recommendations reflect a gender analysis) with an average scoring of 2.41 out of 3. This indicates that integrating GEEW into the scope and evaluation design have improved, more is required to ensure that important findings are also reflected in the conclusions and recommendations.