UNEG
United Nations Evaluation Group

UN-SWAP Evaluation Performance Indicator

2015 Reporting Cycle Results

This document is issued by the UNEG Working Group on Gender Equality and Human Rights and was drafted by Sabrina Evangelista, UN Women Evaluation Specialist and UN-SWAP Evaluation Performance Indicator Focal Point.

April 2016

FOREWORD

The year 2015 was a landmark year for gender equality, as the international community established gender equality as front and central to the achievement of the Sustainable Development Agenda. It is now time for the evaluation community to follow suit.

The United Nations System-wide Action Plan on Gender Equality and the Empowerment of Women (UN-SWAP) is an important accountability mechanism and tool for learning, which has catalyzed real change within UN entities practices with respect to gender equality. As we near the 2017 deadline for compliance with the UN-SWAP, UNEG members are demonstrating their commitment to integrating gender equality in evaluation. Nevertheless, given the current projections for the UN-SWAP Evaluation Performance Indicator, UN entities will not achieve compliance with the Indicator until 2033. This a call for action!

ECOSOC commended UNEG for the action taken in support of the UN-SWAP Framework. The UNEG community must continue its leadership in support of UN-SWAP by making gender-responsive evaluation 'business as usual'.

I look forward to working together to strengthen our evaluation practices and learning from each other with the aim of full compliance with the UN-SWAP by the 2017 deadline.

Marco Segone

UNEG Chair

Table of Contents

Executive Summary

This document has been prepared to inform UNEG members on the 2015 reporting cycle results of the United Nations System-Wide Action Plan on Gender Equality and the Empowerment of Women (UN-SWAP) Evaluation Performance Indicator (EPI).[1] The document provides a brief background of the Chief Executive Board for Coordination (CEB) endorsed UN-SWAP and UNEG support to UN entities to comply with the annual reporting against its Evaluation Performance Indicator (EPI).

The key results from 2015:

1. The vast majority of entities are using the UNEG scorecard[2], however, almost half of those entities that scored "meets requirements" did not use the UNEG Scorecard and thus based their assessment on miscellaneous criteria.
2. Out of the 31[3] entities that used the UNEG scorecard, those with internal reviews were about 4 times more likely than those with an external perspective[4] to score "exceeds" or "meeting" requirements.
3. Although there are only slight differences[5], evaluation reports scored higher on average for inclusion of gender equality in the scope of analysis, indicators and evaluation questions (criteria 1 and 2), and lower regarding use of gender-responsive evaluation methodology/analytical techniques and also inclusion of gender analysis in the findings, conclusions and recommendations (criteria 3 and 4).
4. The aggregate results[6] are pretty evenly split: just over half of UN entities are at least "meeting" requirements, which leaves almost half of UN entities not meeting requirements.
5. Projections show that 100% compliance with the evaluation indicator will not be met until 2033.

The report concludes that concerted and immediate action is needed in order for the UN evaluation community to be in compliance with the UN-SWAP Evaluation Performance Indicator. The report also identifies the need for a deeper analysis of UN-SWAP EPI reporting as a means for better targeting support and facilitating a shared understanding of what it means to meet UNEG Norms and Standards on gender equality. Finally, the report identifies some of the actions evaluation units are taking to ensure gender-responsive evaluation and recommends that UN entities without resources to undertake an external review of the EPI engage in the Peer Learning Exchange in future reporting cycles.

---

[1] This document is issued by the UNEG Working Group on Gender Equality and Human Rights and drafted by Sabrina Evangelista, UN-SWAP Evaluation Performance Indicator Focal Point:
sabrina.evangelista@unwomen.org

[2] See Annex 1 for a full list of the entities, type of review, scores and other details.
[3] This includes GEF, which is not an official UN-SWAP reporting entity
[4] This calculation includes both Peer Learning Exchange and Externally hired companies in the same category.
[5] Not tested for statistical significance.
[6] The aggregate does not include GEF, as it is not an official reporting entity and the aggregate score is reported to ECOSOC.

I.   Background

UN System Wide Action Plan on Gender Equality and the Empowerment of Women

The UN-SWAP constitutes the first accountability framework for gender mainstreaming in the UN system. In Resolution E/RES/2014/12 the Economic and Social Council (ECOSOC) requests the United Nations system, including its agencies, funds and programmes, within their respective organizational mandates, to continue working collaboratively to enhance and accelerate gender mainstreaming within the United Nations system, including by fully implementing the UN-SWAP.[7]

The UN-SWAP is composed of 15 performance indicators for tracking 6 main elements on gender mainstreaming: accountability, results based management, oversight, human and financial resources, capacity, and knowledge exchange and networking. All UN entities are to self-assess and report on progress. UN entities are expected to meet all UN SWAP-performance standards by 2017[8].

Reporting on the UN-SWAP commenced in 2013 and entities are expected to report on a yearly basis through the Report of the Secretary-General to ECOSOC on "Mainstreaming a gender perspective into all policies and programmes in the UN system"[9]. UN Women's Coordination Division provides secretariat services and support to the UN system for reporting on the UN-SWAP. The Gender units of all UN entities play an internal coordinating role, ensuring that progress is accurately reported and that plans of action are developed. Accountability rests, as noted in the CEB policy, with senior managers of the different UN-SWAP reporting entities.

The UN-SWAP framework is accompanied by a set of Technical Notes for each Performance Indicator that provide guidance on how to complete the assessment for each of the 15 Performance Indicators. While the UN-SWAP Performance Indicators approved by the CEB are established, the Technical Notes are considered live documents that can be enhanced.

In 2015, UN Women initiated a review of the implementation of the UN-SWAP, which will result in a revised UN-SWAP framework to be rolled out in 2018.[10]

---

[7] UN-SWAP Framework was developed by UN Women in 2011/2012 in response to the CEB endorsed UN system-wide Policy on Gender Equality and the Empowerment of Women (CEB/2006/2), which was established based on the ECOSOC Agreed Conclusions 1997/2. On 13 April 2012, the CEB endorsed the UN-SWAP for application throughout the UN system. In its resolution E/RES/2012/24 of 27 July 2012, the ECOSOC welcomed the UN-SWAP and called upon the UN system to actively engage in its roll-out and report on the implementation of the resolution at its substantive session in 2013.

[8] There is an extended timeframe to 2019 for those entities with a mainly technical focus

[9] For example see, United Nations, "Mainstreaming a gender perspective into all policies and programmes in the UN system," April 2014; accessible online at:
http://www.un.org/ga/search/view_doc.asp?symbol=E/2014/63&Lang=E

[10] Plans for the roll-out of UN-SWAP/2 are still under consideration, as the framework will need to be endorsed by CEB.

UN-SWAP Evaluation Performance Indicator (EPI)

The oversight element of the UN-SWAP is composed of three performance indicators, including one dedicated to evaluation that is linked to meeting the gender-related UNEG Norms[11] & Standards[12] and demonstrating effective use of the UNEG guidance on integrating gender equality in evaluation[13].

The reporting categories for the Evaluation Performance Indicator (PI5) are as follows:

5. Evaluation Performance Indicator				
Not Applicable	**Missing**	**Approaches**	**Meets**	**Exceeds**
5a. Performance indicator is not relevant to a UN entity	5b. None of the UNEG gender-related norms and standards are met	5c. Meets some of the UNEG gender-related norms and standards	5d. Meets the UNEG gender-related norms and standards	5ei. Meets the UNEG gender-related norms and standards and 5eii. Demonstrates effective use of the UNEG guidance on evaluating from a human rights and gender equality perspective

The UNEG Working Group on Gender Equality and Human Rights developed the Technical Note and Scorecard, which aims to support more systematic and harmonized reporting through the use of a common tool that allows for improved comparability across UN entities. UNEG Heads endorsed the Technical Note and Scorecard in August 2014 (available on the UNEG website: http://www.uneval.org/document/detail/1452). The unit of analysis selected as most feasible was the evaluation report. Thus the Evaluation Performance Indicator should be solely based on an assessment of evaluation reports completed in the reporting year. The Technical Note specifies the below criteria for the assessment of integration of gender equality in the evaluation reports:

1. *GEEW is integrated in the evaluation scope of analysis and evaluation indicators are designed in a way that ensures GEEW-related data will be collected*
2. *GEEW is integrated in evaluation criteria and evaluation questions specifically address how GEEW has been integrated into the design, planning, implementation of the intervention and the results achieved.*
3. *A gender-responsive methodology, methods and tools and data analysis techniques are selected.*

---

[11] United Nations Evaluation Group, Norms for Evaluation in the UN system, 2005; accessible online: http://unevaluation.org/document/detail/21
[12] United Nations Evaluation Group, Standards for Evaluation in the UN system, 2005; accessible online: http://unevaluation.org/document/detail/22
[13] United Nations Evaluation Group, Guidance on Integrating Human Rights and Gender Equality in Evaluation, 2014; accessible online: http://unevaluation.org/document/detail/1616

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

It is recommended that evaluation units conduct an external review, however, as this requires financial resources, the UNEG Working Group on Gender Equality and Human Rights piloted a Peer Learning Exchange (PLE) process in 2015, which proved successful. At a minimum, UN entities are highly encouraged to submit the Scorecard to allow for comparability. Through the UNEG Working Group on Gender Equality and Human Rights, UN Women Independent Evaluation Office provides training webinars on the UN-SWAP EPI reporting process and Scorecard, and help desk support.

II.   UN-SWAP EPI 2015 Reporting Cycle

Methodology

The following results are based on the data that was input for the UN-SWAP Evaluation Performance Indicator by the respective entity UN-SWAP focal point into the online UN-SWAP reporting database[14], which is administered by UN Women. The UN Women EPI focal point accessed the information from the database, accompanying attachments (e.g. the meta-analysis, scorecard, etc.) and remedial actions and liaised with the respective EPI focal point to ensure the information was accurate. It is important to keep in mind that the reporting practices of UNEG members against the UN-SWAP EPI varies greatly, limiting the ability to generalize results across entities or compare with previous reporting cycles.

Evaluation Performance Indicator Results - 2015 Reporting Cycle

For the 2015 reporting cycle 64 entities[15] submitted a report, of these, 18 (N=18/64, 28%) entities reported that the indicator was not applicable to them, which is recommended when the entity does not have an evaluation unit or carry out evaluations as defined by UNEG.[16] The majority (N= 13/18, 72%) of these entities were from the Secretariat. Therefore, these entities are not included in the overall analysis, making the total number of entities reporting that the evaluation indicator is applicable: 46 entities.[17]

---

[14] https://unswap.unwomen.org

[15] The following UN entities did not report against any of the UN-SWAP indicators for the 2015 reporting cycle: UN Interregional Crime and Justice Research Institute, UN Institute for Disarmament Research, UN Research Institute for Social Development; International Civil Aviation Organization; Office of the Special Representative of the Secretary-General for Sexual Violence in Conflict; Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization

[16] The following UN-SWAP reporting entities reported that the Evaluation Indicator was not applicable in 2015: CAAC, DGACM, DM, OAJ, ODA, OHLRLLS, OLA, Ombudsman, OSAA, UNFCCC, UNISDR, United Nations Global Compact, UNOG, UNON, UNOPS, UNSSC, UNU and WMO.

[17] See Annex X for detailed list and results. The GEF is not an official reporting entity results cannot be aggregated to the total results of the UN-SWAP reporting entities but are illustrated with details in a table below.

Use of the UNEG endorsed Scorecard allows for some comparability of results across those entities that use the scorecard. In 2015, the majority of reporting entities used the UNEG Scorecard: 65% (N=30/46)[18], which was also an increase from 2014[19]. Further, in 2015 more entities are conducting external reviews of the integration of gender equality in evaluation reports against the UN-SWAP EPI criteria. In 2015, 22% (N=10/46) of entities submissions were based on an external review: IFAD, ILO, UNDP, UNICEF, UN Women, UNEP and WFP hired external companies/consultants to complete the assessment[20]; and DPI, OHCHR, UNESCO and UNEP[21] participated in the Peer Learning Exchange for the UN-SWAP EPI. The results of external assessments are more objective, allowing for more comparability across entities. Nevertheless, even when entities use the scorecard or conduct an external review, there is much variation according to the number of reports reviewed. Therefore, the aggregate results can be quite misleading. This report presents the disaggregated results first in order to build a better understanding of the aggregate results.

Taking the above points into consideration on comparability, the below analysis is organized as follows: a) comparison between those entities that used UNEG Scorecard; c) comparison between those entities that based their submission on an external review and those that based it on an internal review; and d) aggregate results.

a. The vast majority of entities are using the UNEG scorecard[22], however, almost half of those entities that scored "meets" requirements did not use the UNEG Scorecard and thus based their assessment on miscellaneous criteria.

UNEG encourages the use of the Scorecard in order to have a systematic unit of comparison (the evaluation report), criteria against which to assess and scoring categories corresponding to the UN-SWAP categories. Those evaluation units that did not use the UNEG scorecard based their rating on other factors, such as hiring processes (the extent to which females are hired to conduct evaluations) or systems in place to support gender equality in evaluation (such as an organizational policy on gender equality). The UNEG scorecard allows for a minimum level of coherence in reporting.

---

[18] Although DESA submitted a scorecard, it was not in compliance with the Technical Guidance as it was not based on evaluation reports.

[19] In the 2014 reporting cycle only 3 entities conducted an external assessment: UNDP, UN Women and WFP.

[20] It should be noted that both ILO and UNICEF submissions were based on review of evaluation reports completed in 2014 (and in first quarter of 2015 in the case of ILO) due to the misalignment of reporting processes. It was decided by UNEG Heads that an external review could be based on evaluation reports completed in the previous year. Efforts to align meta-evaluation processes with the UN-SWAP reporting cycle are encouraged.

[21] UNEP hired an external consultant and participated in the Peer Learning Exchange, but is not counted twice.

[22] See Annex 1 for a full list of the entities, type of review, scores and other details.

In 2015, the vast majority of entities used the UNEG scorecard and Technical Note (66%, N=31/47).[23] However, it is important to note that almost half (43%, N=9/21) of those entities that scored "meets requirements" in 2015, did not use the UNEG scorecard[24] (see Figure 1, Table 1 or Annex 1).

Figure 1. Number of entities per total in reporting category by UNEG scorecard usage (N=47)
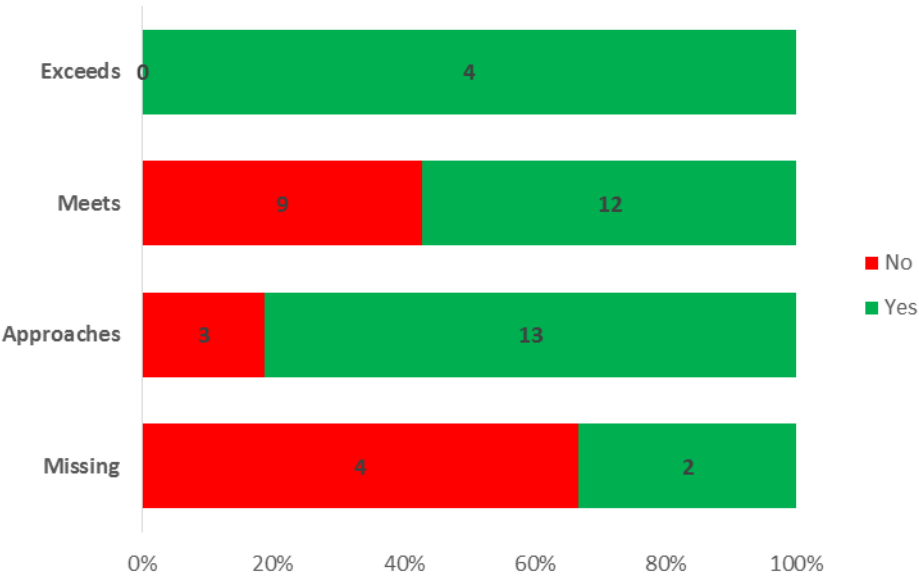
b. **Out of the 31[25] entities that used the UNEG scorecard, those with internal reviews were about 4 times more likely than those with an external perspective[26] to score exceeds or meeting requirements.**

It is important to examine differences, if any, between those entities that used the scorecard with an internal review vs. those that had an external perspective. As noted above, 10 submissions were based on an external review: IFAD, ILO[27], UNDP, UNICEF[28], UN Women,

---

[23] This figure includes GEF, however, GEF was not an official UN-SWAP reporting entity in 2015. The list of entities that used the Scorecard is found in Annex 1.

[24] The following entities did **not** use the UNEG scorecard and Technical Note in 2015: ECA; ITU; PBSO; DESA[24]; DFS; DPKO; IAEA; IMO; ITC-ILO; UNAIDS; UNHCR; UNIDO[24]; DPA; UNITAR; UNWTO; and UPU.

[25] This includes GEF, which is not an official UN-SWAP reporting entity

[26] This calculation includes both Peer Learning Exchange and Externally hired companies in the same category.

[27] It should be noted that both ILO and UNICEF submissions were based on review of evaluation reports completed in 2014 (and in first quarter of 2015 in the case of ILO) due to the misalignment of reporting processes. It was endorsed on a non-objection basis by UNEG Heads that an external review could be based on evaluation reports completed in the previous year. Efforts to align meta-evaluation processes with the UN-SWAP reporting cycle are encouraged.

[28] It should be noted that both ILO and UNICEF submissions were based on review of evaluation reports completed in 2014 (and in first quarter of 2015 in the case of ILO) due to the misalignment of reporting processes. It was endorsed on a non-objection basis by UNEG Heads that an external review could be based on evaluation reports completed in the previous year. Efforts to align meta-evaluation processes with the UN-SWAP reporting cycle are encouraged.

UNEP and WFP hired external companies/consultants to complete the assessment; and DPI, OHCHR, DPI and UNEP[29] participated in the Peer Learning Exchange for the UN-SWAP EPI.

As the overall UN-SWAP was built as a self-assessment tool, it is important to note that the Evaluation Indicator is the only category that advocates for an external assessment because external assessments are deemed to be more objective than internal/self-assessments. All other categories of the UN-SWAP are assessed internally. There are definite pros to having an internal assessment, particularly in the early years of UN-SWAP implementation, as one colleague noted, "The internal assessment process helped to engage evaluation office staff in really understanding the issues and exploring and committing to action to improve the rating."[30]

However, internal assessments are also more likely to exhibit bias, and the results show that **entities with internal reviews were almost 4 times more likely to report a higher score (exceeds/meets) than entities that had an external review.[31]** None of the 4 entities that scored "exceeds requirements" hired an external reviewer – although one (OHCHR) participated in the peer learning exchange. Figure 2 illustrates how internal reports were skewed towards the positive as compared with Peer Learning Exchange and externally reviewed reports.

Figure 2. Number of entities per reporting category and type of review (N = 31)

---
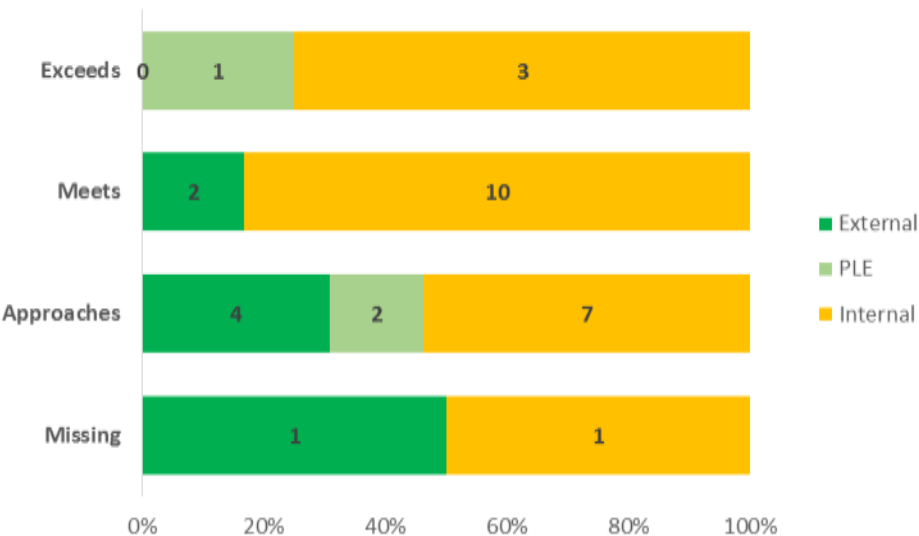
[29] UNEP hired an external consultant and participated in the Peer Learning Exchange, but is not counted twice.
[30] Feedback from UNEP UN-SWAP EPI Focal Point.
[31] For the analysis, the Peer Learning Exchange was lumped together with the External Reviews; however, the results are disaggregated by category in the Table. Due to the low numbers, statistical significance is not possible; the odds of scoring meets/exceeds were calculated and then an odds ratio was calculated: 13/8= Internal reports are 1.6 times more likely to score exceeds/meets than approaches/missing; 3/7 = External reports are .42 times more likely to score exceeds/meets than approaches/missing; 1.6/.42 = 3.8 Internal reports are 3.8 times more likely to score exceeds/meets than external reports.

As can be seen from Table 1 below, out of the 4 entities that "exceed requirements" only 1 (25%) participated in the peer learning exchange and none had an external review; and they all had 10 reports or fewer. Out of the 12 entities that "meet requirements", only 2 (17%) entities had an external review. Whereas, out of the 13 entities that reported, "approaches requirements" 6 (46%) were based on external assessments.

An important point is that there are also differences in the types and quantity of reports assessed. Many of those externally reviewed who scored "approaches" have decentralized evaluation functions, whereas the majority of those that scored "Meets/Exceeds" do not have a decentralized evaluation function.[32] It may be the case that the decentralized function fares worse than the centralized function in terms of gender-responsiveness due to the lack of overall M&E skillset at the decentralized level. Additionally, given that the majority of the externally reviewed reports were looking at 21 or more reports, the higher quantity of reports may also provide more robustness to the assessment and may reflect that those entities have more resources to be able to carry out such an external review.

| Table 1. Disaggregated results for UN entities using the UNEG Scorecard (N=31)					
# Reports → Rating ↓	1-2 Reports	3-5 Reports	6-10 Reports	11-15 Reports	21 or more reports
Exceeds	UNRWA	OHCHR^	OIOS ESCWA		
Meets	DSS OCHA UNCDF	ESCAP ITC WHO WIPO	ECLAC	UNFPA	FAO IFAD** WFP**
Approaches	DPI^	UN Habitat UNV	GEF[1] IOM UNCTAD	UNESCO^ UNODC UNOV	UNDP** ILO** UN Women** UNICEF**
Missing	ECE				UNEP^**
^Participated in PLE **External Review [1]GEF is included in this table and section, but not in the overall aggregate figures reported in the report, as they are not an official reporting entity to the UN-SWAP in 2015					

---

[32] For example, UNDP, UN Women, UNICEF and UNESCO included evaluation reports from the decentralized function; and UN Women did not complete any corporate evaluation in 2015 and thus the score is entirely based on decentralized evaluations. On the other hand, IFAD and FAO (for example) only included corporate evaluations.

c. **Although there are only slight differences[33], evaluation reports scored higher on average for inclusion of gender equality in the scope of analysis, indicators and evaluation questions (criteria 1 and 2), and lower regarding use of gender-responsive evaluation methodology/analytical techniques and also inclusion of gender analysis in the findings, conclusions and recommendations (criteria 3 and 4).**

The UNEG scorecard allows for a closer look at how each evaluation report is faring across the criteria. The analysis shows that out of the 389 reports reviewed by UN entities, on average, evaluations were scoring higher for inclusion of gender equality in the scope of analysis, indicators and evaluation questions (criteria 1 and 2), and lower regarding use of gender-responsive evaluation methodology/analytical techniques and also inclusion

UNEG EPI Scorecard Scoring Criteria	
0	not at all integrated
1	partially integrated
2	satisfactorily integrated
3	fully integrated

of gender analysis in the findings, conclusions and recommendations (criteria 3 and 4). The analysis also shows that internally reviewed reports scored on average higher across all indicators compared with the externally reviewed reports.

Figure 3. Average score by criteria and type of review[34]

---
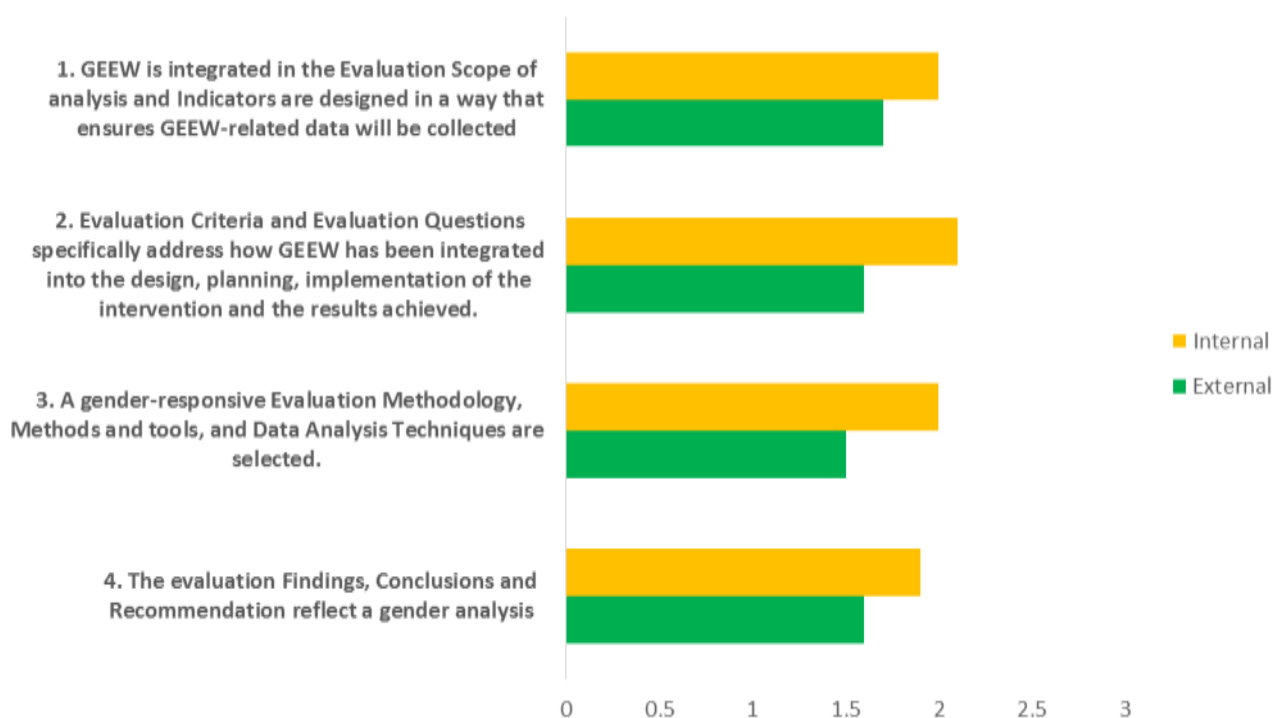
[33] Not tested for statistical significance.
[34] These figures include GEF. External and Peer Learning Exchange were combined.

d. The aggregate results[35] are pretty evenly split: just over half of UN entities are at least meeting requirements, which leaves almost half of UN entities not meeting requirements.

It is important to keep in mind the above-disaggregated analysis and aforementioned limitations with comparability when looking at the aggregate results (which includes both entities that did use the UNEG Scorecard and those that did not), as the aggregate masks within category differences in reporting. However, because the aggregate scores are submitted to ECOSOC[36], these are presented in this report. When we look at the results across all entities that reported the indicator is applicable (N=46), just over half of UN entities (N=25, 54%) are meeting or exceeding the requirements for the EPI.[37] However, there are still almost half (N= 21, 46%) that are only approaching or missing requirements. It is important to note that of those entities that are "meeting" requirements, almost half (see section a above) of them did not use the UNEG scorecard and thus are basing their assessment on different factors than the majority of other entities.

Figure 4. Aggregate Evaluation Performance Indicator Rating, 2015 Reporting Cycle (N=46)

When we look at the breakdown by type of entity, the ratings are also spread pretty evenly, with the exception of the Training Institute[38].

Figure 5. Rating by type of entity (for both entities that did/not use UNEG Scorecard)

---

[35] The aggregate does not include GEF, as it is not an official reporting entity and the aggregate score is reported to ECOSOC.

[36] The Annual Secretary General's report on Gender Mainstreaming in the UN system only looks at the aggregate UN-SWAP indicator results and not within indicator differences.

[37] DPKO/DFS and UNODC/UNOV submit only one report but it counts double for all UN-SWAP performance indicators.

[38] UNITAR, all other training institutes reported N/A.

e. Projections show that 100% compliance with the evaluation indicator will not be met until 2033.

The UN Women Coordination Division calculated projections for each of the UN-SWAP indicators based on the average annual per cent change in rating from 2012-2015 (See Annex 2 for the detailed list of ratings for 2014 and 2015 when the same UNEG criteria were applied). When all reporting entities are included in the calculation the results are stark: compliance with the evaluation indicator will not be met until 2033; this figure includes those entities that reported the EPI was "not applicable" and thus may not currently have an evaluation function[39]. However, even when we exclude those entities[40], compliance is still far off: 2022 (7 years from today).

Table 2. Projected year to attain 100% compliance with UN-SWAP EPI

Indicator	A = Avg Annual % change 2012-2015	B = 2015 Meets & Exceeds (% of total ratings)	C = Projected number of years to attainment (100-B)/A	Projected Year to attain 100% compliance
PI 5: Evaluation – including N/A*	3.3	39	**18**	**2033**
PI 5: Evaluation – excluding N/A*	6.3	54	**7**	**2022**
*N/A = is most typically reported by those entities that do not have an evaluation function				

Conclusions

The results of the 2015 reporting cycle indicate that UNEG members are making progress integrating gender equality in evaluation reports; however, there is much more work to be done. If the UN evaluation community is to be in compliance with UN-SWAP, it must take action quickly, otherwise it will be lagging far behind the expected deadline.

The results also demonstrate that although the use of the UNEG scorecard has contributed to more coherent reporting against the Evaluation Performance Indicator through the use of a unit of analysis, the differences between internal and external reviews may indicate that there are different interpretations on what it means to integrate gender equality in evaluation

---

[39] This figure was calculated based on extrapolation of the average annual percentage change; also please note that due to rounding some estimates are actually slightly past the years shown.
[40] The number excluding N/A may be more realistic, as the number reporting N/A is a relatively static proportion of reporting entities since those that do not have an evaluation function will not typically change reporting status.

reports and meet UNEG Norms and Standards. Given the average scores on the criteria, the results also indicate that concerted effort needs to be made to ensure that not only the questions and scope of analysis are gender-responsive, but that this carries through to the evaluation findings and conclusions.

III.   Way forward

Given the current projections for compliance with the UN-SWAP Evaluation performance indicator, the evaluation community needs to make a concerted push to place gender equality front and central in its evaluation practice in order to meet requirements by the 2017 deadline.

The current figures demonstrate that it is time for evaluation units to take stock and analyse the underlying bottlenecks and challenges to gender-responsive evaluation. Each entity is unique and operates within a different environment and with a different technical focus, thus there is no "one size fits all" solution. However, given that we now have 4 reporting years behind us and a wealth of information from the entities that used scorecards, it is time to dig deeper to look beneath the numbers and develop an understanding of how the Evaluation Performance Indicator and scorecard criteria are being applied i.e. rationale for the scoring. This will enable us to identify where there may be gaps in knowledge or misinterpretation and help us to facilitate a shared understanding of what it means to meet UNEG Norms and Standards on gender equality. This will also help to identify good practice and the bottlenecks to integrating gender equality, and to better target support to UN entities. Given that the entire UN-SWAP framework will be revised in 2016 (for implementation in 2018), it is an opportune moment to analyse these underlying issues and assess whether any changes should be made to the Evaluation Performance Indicator. Furthermore, the UNEG Peer Learning Exchange for the UN-SWAP EPI, which was piloted for the 2015 reporting cycle, is encouraged for those entities that are unable to carry out an external review.

It is important to acknowledge some of the remedial actions UN entities are taking to strengthen the gender-responsiveness of evaluation systems and practice, as we can learn from each other:

- Revising the evaluation policy and guidelines to better integrate gender equality and refer to UNEG Guidance on Human Rights and Gender Equality
- Establishing quality assurance mechanisms (both internal and external) that review integration of gender equality throughout the evaluation process
- Organizing joint training on gender-responsive evaluation
- Developing a template for Terms of Reference that specify sample questions that can be adapted as appropriate
- Ensuring gender expertise on evaluation teams

- Establishing Professionalization initiatives to support strengthening of gender-responsive evaluation skills of decentralized evaluation

There are also many existing resources and resources under development (to be launched soon) to support UN entities in ensuring a gender-responsive evaluation:

- In 2016, UN Women launched the Gender Evaluation Consultant Database of consultants with gender and evaluation expertise: http://evalconsultants.unwomen.org
- In 2015, UN Women launched an eLearning course: How to manage gender-responsive evaluation – mainly targeted at UN Women staff, but also relevant to others: http://trainingcentre.unwomen.org
- UNEG is developing an eLearning course to complement the UNEG Guidance on Human Rights and Gender Equality (to be launched in 2016)
- UNEG is developing a framework for Professionalization
- UNEG Norms and Standards are under revision
- The UNEG Working Group on Gender Equality facilitates development of guidance, sharing of knowledge on integration human rights and gender equality through webinars (available on the UNEG Youtube) and "In Focus" series, which highlights how UN evaluation units are changing their practices to become gender-responsive, and supports the UN-SWAP EPI Peer Learning Exchange[41].

---

[41] If you are interested in participating in the Peer Learning Exchange for the 2016 reporting cycle, contact: Sabrina.evangelista@unwomen.org

Annex 1

UN entities that used the UNEG Scorecard

Type of Review	Entity short name	Type of Entity	Rating	Avg Criteria 1	Avg Criteria 2	Avg Criteria 3	Avg Criteria 4	Overall Score	#reports reviewed	Range of reports
External	IFAD	Funds & Programmes	Meets	2.8	2.6	2.1	2.7	10.2	21	equal to 21 or more
External	UNICEF	Funds & Programmes	Approaches	1.64	1.52	1.42	1.42	5.74	69	equal to 21 or more
External	UNEP	Secretariat	Missing	1.03	0.8	0.43	0.63	2.9	35	equal to 21 or more
External	ILO	Specialized	Approaches	1.2	1.1	1	1.3	4.6	42	equal to 21 or more
External	UNDP	Funds & Programmes	Approaches	1.76	1.24	1.4	1.52	6	25	equal to 21 or more
External	WFP	Funds & Programmes	Meets	2.14	1.95	2.24	1.76	8.19	21	equal to 21 or more
External	UN Women	Funds & Programmes	Approaches	1.9	1.9	1.6	1.2	6.6	27	equal to 21 or more
PLE	UNESCO	Specialized	Approaches	0.8	0.87	0.87	1.27	3.8	15	equal to 11 - 15
PLE	DPI	Secretariat	Approaches	1	1.5	1.5	2	6	2	equal to 1-2
PLE	OHCHR	Secretariat	Exceeds	3	3	2.3	2.3	10.66	3	equal to 3-5
Internal	FAO	Specialized	Meets	2.1	2.3	2.4	1.7	8.54	22	equal to 21 or more
Internal	IOM	Funds & Programmes	Approaches	1.5	1.4	1.5	0.88	5.25	8	equal to 6 - 10

Internal	OCHA	Secretariat	Meets	3	2	2	2	9	2	equal to 1-2
Internal	ECE	Secretariat	Missing	0	0	1	1	2	2	equal to 1-2
Internal	ECLAC	Secretariat	Meets	2.75	2.75	2.75	1.875	10.125	8	equal to 6 - 10
Internal	UNCTAD	Secretariat	Approaches	2	1.7	2.1	1.6	7.43	7	equal to 6 - 10
Internal	ESCWA	Secretariat	Exceeds	3	2.83	3	3	11.75	6	equal to 6 - 10
Internal	UN Habitat	Secretariat	Approaches	1.33	1.67	1.67	1.67	6.33	3	equal to 3-5
Internal	UNFPA	Funds & Programmes	Meets	2.3	2	2.1	2.4	8.87	15	equal to 11 - 15
Internal	UNODC	Secretariat	Approaches	0.9	1.5	0.8	1.2	4.36	11	equal to 11 - 15
Internal	UNV	Funds & Programmes	Approaches	2	1.67	1	1	5.67	3	equal to 3-5
Internal	ITC	Funds & Programmes	Meets	2.2	2.2	3	3	10.4	5	equal to 3-5
Internal	UNCDF	Funds & Programmes	Meets	2	2.5	2.5	2.5	9	2	equal to 1-2
Internal	UNRWA	Funds & Programmes	Exceeds	3	3	3	3	12	1	equal to 1-2
Internal	DSS	Secretariat	Meets	2	3	2	3	10	1	equal to 1-2
Internal	ESCAP	Secretariat	Meets	2	3	2	1.75	8.75	4	equal to 3-5
Internal	OIOS	Secretariat	Exceeds	2.83	2.83	3	2.17	10.83	6	equal to 6 - 10
Internal	UNOV	Secretariat	Approaches	0.9	1.5	0.8	1.2	4.36	11	equal to 11 - 15

Internal	WHO	Specialized	Meets	2.67	2.67	2.33	1.67	9.33	3	equal to 3-5
Internal	WIPO	Technical	Meets	2	2	2.33	2.33	8.67	3	equal to 3-5
Internal	GEF (not an official reporting entity)	Funds & Programmes	Approaches	0.83	0.67	1	1.5	4	6	equal to 6 - 10

Annex 2

Ratings in 2014 & 2015 for those entities that used the UNEG Scorecard

Entity short name	2014			2015			Change in rating bt years
	Rating	Score	# reports	Rating	Score	#reports	
IFAD*	Meets	9.66	15	Meets	10.2	21	SAME
UNICEF*	Approaches	6	15	Approaches	5.74	69	SAME
UNEP*	Missing	2.2	17	Missing	2.9	35	SAME
ILO*	Did not use scorecard			Approaches	4.6	42	
UNDP*	Approaches	5.9	30	Approaches	6	25	SAME
WFP*	Meets	7.26	19	Meets	8.19	21	SAME
UN Women*	Meets	7.57	22	Approaches	6.6	27	Decrease
UNESCO*	Approaches	3.53	19	Approaches	3.8	15	SAME
DPI*	Meets	9	1	Approaches	6	2	Decrease
OHCHR*	Meets	7.66	3	Exceeds	10.66	3	Increase
FAO	Meets	7.68	28	Meets	8.54	22	SAME
IOM	Approaches	4.73	11	Approaches	5.25	8	SAME
OCHA	Meets	9	1	Meets	9	2	SAME
ECE	Missing	2	7	Missing	2	2	SAME
ECLAC	Approaches	7.25	4	Meets	10.125	8	Increase
UNCTAD	Approaches	6.71	7	Approaches	7.43	7	SAME
ESCWA	Meets	8.33	3	Exceeds	11.75	6	Increase
UN Habitat	Approaches	7	6	Approaches	6.33	3	SAME
UNFPA	Meets	9	11	Meets	8.87	15	SAME
UNODC	Approaches	4	6	Approaches	4.36	11	SAME
UNV	Approaches	6	2	Approaches	5.67	3	SAME
ITC	Did not use scorecard			Meets	10.4	5	

UNCDF	Meets	7.66	1	Meets	9	2	SAME
UNRWA	Approaches	6.25	6	Exceeds	12	1	Increase
DSS	Did not use scorecard			Meets	10	1	
ESCAP	Did not use scorecard			Meets	8.75	4	
OIOS	Exceeds	10.85	6	Exceeds	10.83	6	SAME
UNOV	Approaches	4	6	Approaches	4.36	11	SAME
WHO	Meets	4.5	2	Meets	9.33	3	SAME
WIPO	Approaches	4.67	3	Meets	8.67	3	Increase
UNIDO	Meets requirements	7.78	18	Did not use scorecard			
PBSO	Approaches	6	3	Did not use scorecard			
GEF (not an official reporting entity)	Missing	2.6	5	Approaches	4	6	Increase

* Indicates External Assessment or Peer Learning Exchange