



UNEG
United Nations Evaluation Group

Working
Paper

UN-SWAP Evaluation Performance Indicator 2020 Reporting Cycle Results

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

June 2021

FOREWORD

This report presents results, trends and insights from the 2020 reporting cycle for the UN-SWAP Evaluation Performance Indicator (EPI). The UN-SWAP EPI aims to capture the gender responsiveness and effective use of UNEG guidance on integrating human rights and gender equality, during all phases of evaluations conducted by various reporting entities.

The UN-SWAP 2.0 was introduced in 2018, and 2020 was the third year of reporting on the revised guidelines. Following the trend set in previous years, the number of reporting entities meeting or exceeding UN-SWAP EPI requirements in 2020 has increased. There is also continued adoption of the UNEG-endorsed scorecard for reporting. The use of the scorecard is essential to enabling cross-entity comparisons. Another good practice is the use of independent or peer-reviewed assessments, and we hope that more entities will be encouraged to use either of these modalities in future reporting cycles. Overall, the quality of EPI reporting is improving, with many entities providing examples of innovative efforts to support gender mainstreaming in evaluation.

The COVID-19 pandemic led to delays and disruptions in the evaluation activities of several entities and this is likely to continue in 2021. Entities have made efforts to adapt to the ongoing pandemic, particularly by improving methods for virtual data collection and issuing guidance documents on conducting remote evaluations. This report provides several examples of these rich resources, that are available for all entities to draw on and learn from. Evaluation units have also produced guidance documents for conducting decentralized evaluations and rapid assessments, among others.

UN Women Independent Evaluation Service (IES) serves as the secretariat for the UN-SWAP EPI and through this annual report, we present progress, good practices and areas of potential improvement. As we conduct evaluations in these challenging times, we hope this report is a helpful resource in supporting our goal of advancing gender mainstreaming in evaluations.



Inga Sniukaite

Chief of Independent Evaluation Services

UN Women

Table of Contents

FOREWORD	i
Acronyms	1
Executive Summary.....	2
1. Background	3
United Nations System-Wide Action Plan on Gender Equality and the Empowerment of Women (UN-SWAP)	3
UN-SWAP Evaluation Performance Indicator (EPI).....	4
Evaluation Performance Indicator Methodology	4
2. Evaluation Performance Indicator Results.....	7
Key Insights	7
Insight 1: There is a consistent increase in the proportion of reporting entities meeting or exceeding UN-SWAP EPI requirements in 2020.....	7
Insight 2: There is continued adoption of the UNEG-endorsed scorecard for reporting on the EPI, and a substantial improvement in the overall ratings of scorecard using entities.	8
Insight 3: There is an increase in the number of evaluations included in the EPI assessment sample and an overall improvement in the quality of reporting.	10
Insight 4: COVID-19 related disruptions were indicated by nearly three fourths of reporting entities and are likely to continue to affect evaluation activities in 2021.....	11
Good practices to advance integration of gender equality in evaluations.....	12
3. Way forward	14

Acronyms

EPI	Evaluation Performance Indicator
GEWE	Gender Equality and Women's Empowerment
IEAS	UN Women Independent Evaluation and Audit Services
IES	Independent Evaluation Service
OECD -DAC	Organisation for Economic Co-operation and Development - Development Assistance Committee
SDG	Sustainable Development Goals
UNEG	United Nations Evaluation Group
UN-SWAP	United Nations System-Wide Action Plan on Gender Equality and the Empowerment of Women

Executive Summary

The UN-SWAP Evaluation Performance Indicator (EPI) assesses the extent to which evaluation reports of an entity meet the gender-related United Nations Evaluation Group (UNEG) Norms and Standards and demonstrate effective use of the UNEG Guidance on integrating human rights and gender equality during all phases of the evaluation.

This report presents insights from the 2020 reporting cycle, based on reports and documents submitted to the online UN-SWAP database by reporting entities. The key insights from the 2020 reports are as follows:

Insight 1: There is a consistent increase in the proportion of reporting entities meeting or exceeding UN-SWAP EPI requirements in 2020.

Insight 2: There is continued adoption of the UNEG-endorsed scorecard for reporting on the EPI, and a substantial improvement in the overall ratings of scorecard using entities.

Insight 3: There is an increase in the number of evaluations included in the EPI assessment sample and an overall improvement in the quality of reporting.

Insight 4: COVID-19 related disruptions were indicated by nearly three fourths of reporting entities and are likely to continue to affect evaluation activities in 2021.

2020 was the third year of reporting on the revised UN-SWAP 2.0. The quality of the EPI reporting continued to be high, with detailed reports highlighting good practices and innovative efforts to support gender mainstreaming in evaluation. Many of these were related efforts to adapt evaluation methods to the ongoing global pandemic. Entities have also reported integrating UN-SWAP indicators into their internal quality assessment frameworks, which will serve to further enhance the quality of gender mainstreaming in evaluations. While an increasing number of entities are using the UNEG endorsed scorecard, more entities should commission independent, or peer-review based assessments in the future.

The COVID-19 pandemic will continue to impact evaluations conducted in 2021. While most agencies reported disruptions owing to the pandemic, many have adjusted evaluation methods, issued guidance documents, and suggested ways to undertake and improve the quality of remote evaluations. There is a good opportunity for entities to learn from each other's experiences and use the newly developed resources on conducting remote evaluations.

As in previous years, evaluation units of various entities have produced knowledge products on different aspects of gender responsive evaluations, including guidance documents for decentralized evaluations, pocket tools, rapid assessments, and gender frameworks for evaluations.

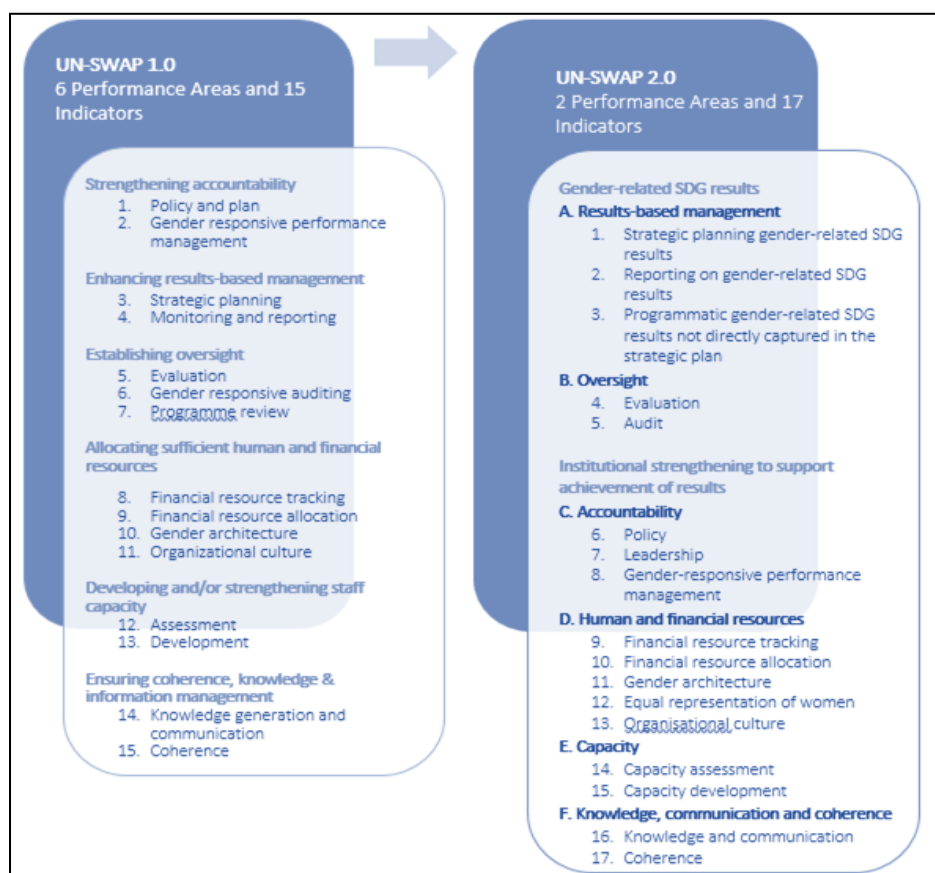
In 2021, the UNEG working group on Gender, Disability and Human Rights plans to undertake a gap analysis of key guidance documents, including the UNEG guide *on Integrating Human Rights and Gender Equality in Evaluations* to identify potential improvements such as disability inclusion and the new Organisation for Economic Co-operation and Development Development Assistance Committee(OECD DAC) evaluation criteria.

1. Background

United Nations System-Wide Action Plan on Gender Equality and the Empowerment of Women (UN-SWAP)

UN entities annual report progress on the UN-SWAP accountability framework on gender equality and the empowerment of women. In 2018, a revised UN SWAP 2.0 was launched and included improvements in the guidance documents for each of the 17 performance indicators (see Figure 1). The technical note for the Evaluation Performance Indicator¹ (EPI) was revised to improve coherence in reporting across entities and minimize subjectivity in applying scoring criteria. 2020 is the third year of implementation for the UN-SWAP 2.0. Based on feedback received from reporting entities, the EPI technical note was further updated in 2020 with the support of the UNEG working group on Gender, Disability and Human Rights to enhance the clarity of assessment criteria.

Figure 1: Transition from UN-SWAP 1.0 → UN-SWAP2.0



¹ To view the revised technical note, please download at: <http://www.uneval.org/document/download/2148>

UN-SWAP Evaluation Performance Indicator (EPI)

The UN-SWAP EPI assesses the extent to which evaluation reports of an entity meet the gender-related United Nations Evaluation Group (UNEG) Norms and Standards² and demonstrate effective use of the UNEG Guidance on integrating human rights and gender equality during all phases of the evaluation.

The UN-SWAP EPI technical note and scorecard establish guidance and a minimum set of criteria to capture the overall elements related to mainstreaming gender equality in evaluation. The requirements are aligned with UNEG norms, standards, and guidance on how to integrate gender and human rights into evaluations. The technical note also encourages all reporting UN system entities to conduct at least one evaluation to assess corporate performance on gender mainstreaming every 5-8 years. This might constitute, but is not limited to, corporate evaluations of gender policy, mainstreaming, and strategy³.

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’.

Evaluation Performance Indicator Methodology

In line with other UN-SWAP indicators, the EPI is linked to a five-level rating system, with the following categories: “not applicable”, “misses requirements”, “approaches requirements”, “meets requirements”, and “exceeds requirements”. The three reporting criteria for the EPI are as follows:

- **Approaches requirements**
4a. Meets some of the UNEG gender-related norms and standards in the UNEG guidance on Integrating Human Rights and Gender Equality in Evaluation

- **Meets requirements**
4bi. Meets the UNEG gender-related norms and standards
and
4bii. Applies the UNEG guidance on Integrating Human Rights and Gender Equality in evaluation during all phases of the evaluation

- **Exceeds requirements**
4ci. Meets the UNEG gender-related norms and standards
and

² To view a full list of UNEG Norms and Standards, please download at: <http://www.uneval.org/document/detail/1914>

³ UN Women IES provides help desk services and support to the UN system for reporting on the UN-SWAP EPI.

4cii. Applies the 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 corporate performance on gender mainstreaming or evaluation of its gender equality policy/strategy every 5-8 years

An entity is expected to report “not-applicable” if there is no evaluation unit and no evaluations are conducted by the entity. In case an entity has conducted evaluations previously, but not in the reporting year, the last rating completed should be used with a clear note indicating the year upon which the rating is based. This approach avoids confusion with those entities that do not have an evaluation unit.

In order to assess overall progress against the criteria, entities undertake an assessment of individual evaluations. Entities are advised to employ the accompanying scorecard⁴ and guiding questions mentioned below in Table 1. The use of the scorecard ensures a harmonized reporting across entities and a more rigorous EPI assessment. Thus, the EPI is primarily based on an assessment of evaluation reports completed in the reporting year.

Table 1. UN-SWAP EPI criteria for assessing evaluation reports⁵

Scoring Criteria		Guiding questions for assessing integration
Criterion 1	GEWE is integrated in the evaluation scope of analysis and evaluation criteria and questions are designed in a way that ensures GEWE-related data will be collected.	<ul style="list-style-type: none"> a. Do the evaluation objectives and/or scope include analysis of the extent to which HR&GE were taken into consideration in the design of the programme/project/policy being evaluated and the achievement of HR&GE-related results? b. Does the evaluation assess whether sufficient information was collected during the implementation period on specific indicators to measure progress on HR&GE? c. Was a stand-alone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria by being gender-disaggregated, gender-specific (relevant to a specific social group), or gender-focused (concerning relations between social groups)? d. Is there a dedicated evaluation question regarding how GEWE has been integrated into the design, planning and implementation of the intervention and the results achieved or integrated throughout other questions?
Criterion 2	Gender-responsive methodology, methods, tools, and data analysis techniques are selected.	<ul style="list-style-type: none"> a. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating HR&GE considerations? Are a diverse range of data sources and processes employed (i.e., triangulation, validation)? Was data disaggregated by sex? b. Were methods used for ensuring meaningful participation and the inclusion of women’s voices as well as underrepresented

⁴ To view the scorecard, please download at: <http://www.unevaluation.org/document/download/2149>

⁵ The first three criteria are based on an assessment of evaluation reports.

		<p>groups, including the most vulnerable where appropriate, throughout the evaluation process (inception, data collection and reporting phases)?</p> <p>c. Does the sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable?</p> <p>d. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality?</p>
Criterion 3	Evaluation findings, conclusions and recommendations reflect a gender analysis.	<p>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue that is being addressed by the evaluation?</p> <p>b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different groups, and/or disaggregates quantitative data?</p> <p>c. Are unanticipated effects of the intervention on HR&GE described?</p> <p>d. Do the findings, conclusions and recommendations explicitly address the gender and human rights dimensions assessed by the evaluation?</p> <p>e. Does the evaluation report provide specific recommendations addressing GEWE issues and priorities for action to improve GEWE of the intervention or future initiatives in this area?</p>
Criterion 4	At least one evaluation to assess corporate performance on gender mainstreaming is conducted every five to eight years.	In order to “exceed requirements”, an evaluation report’s average score must “meet requirements” and the entity must also conduct an evaluation of its corporate gender policy or equivalent.

Three modes of assessment for the UN-SWAP EPI are recommended in the technical note. These include self-assessment, peer-review, and external review conducted by an independent consultant. In 2020, nine entities reported conducting an external/independent assessment, two engaged in peer-reviews, and the remaining opted for self-assessment. For the purpose of the assessment, entities are advised to include a representative sample of evaluation reports. Some entities chose to include the entire universe of evaluations, while others include a sample of corporate and decentralized evaluations⁶. The samples draw on different thematic and geographic areas to provide appropriate coverage.

⁶ The number of evaluations included in the reporting sample ranged from 1 to 112 in 2020.

2. Evaluation Performance Indicator Results

Key Insights

Insight 1: There is a consistent increase in the proportion of reporting entities meeting or exceeding UN-SWAP EPI requirements in 2020.

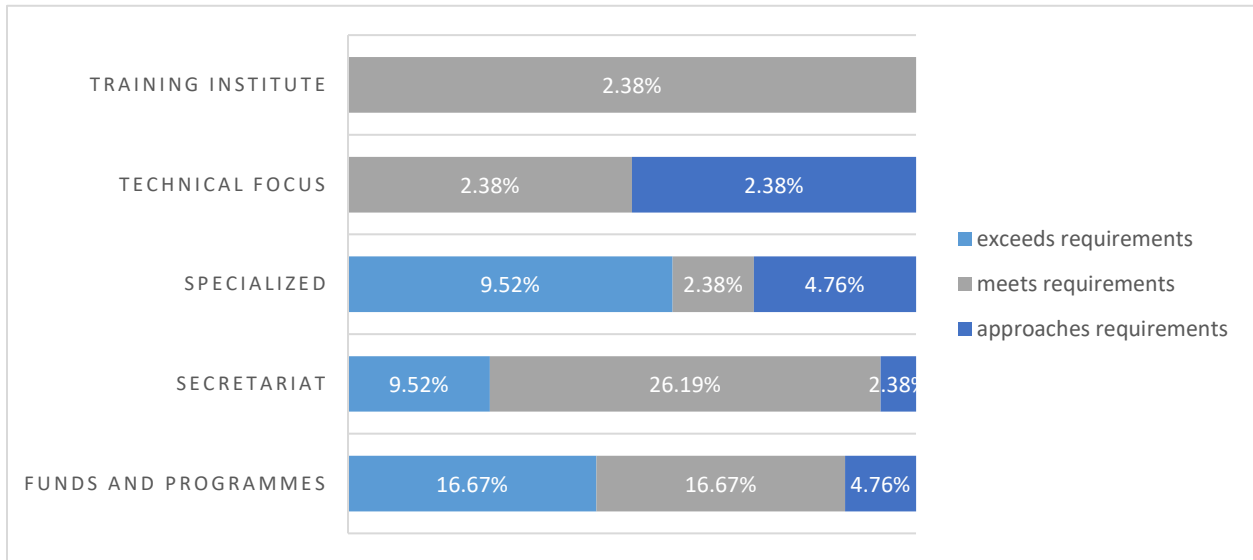
In the 2020 reporting cycle, 42 out of 66 entities reported progress on the UN-SWAP EPI, while 24 entities submitted a rating of “not applicable”. Table 2 presents a disaggregation of ratings by entity type.

Table 2. Disaggregated results for UN-SWAP entities in 2020: by EPI rating and entity type (N=66)

		EPI Rating			
		Exceeds requirements	Meets requirements	Approaches requirements	Not Applicable
Entity Type	UN Secretariat	4	11	1	18
	Funds and Programmes	7	7	2	1
	Specialized Agencies	4	1	2	1
	Technical Focus	-	1	1	2
	Training Institute	-	1	-	2
	Grand Total	15	21	6	24

Of the 42 entities reporting against the EPI, 83 per cent reported meeting or exceeding requirements in 2020. This represents an increase of 13 percentage points compared with 2019. 36 per cent of all reporting entities are now exceeding requirements, which is an increase of four percentage points compared with 2019 and in line with the historical (upward) trend in this indicator. Figure 2 presents the results by entity type.

Figure 2. Disaggregated results (percentage) for UN-SWAP EPI reporting entities in 2020: by EPI rating and entity type (N=42)



Insight 2: There is continued adoption of the UNEG-endorsed scorecard for reporting on the EPI, and a substantial improvement in the overall ratings of scorecard using entities.

Compared with 2019, there was an increase in the number of reporting entities that used the UNEG-endorsed scorecard to assess performance for the EPI. This is particularly encouraging as the proportion of scorecard users had slightly declined in 2019. The scorecard enables a harmonized approach in assessing the comparative performance of entities that vary by size, mandate, and capacity. In an effort to improve and encourage the use of the scorecard, the EPI technical note was further updated in 2020 by the UNEG working group on Gender, Disabilities and Human Rights to improve the clarity of guidance. The entities that did not use the UNEG-endorsed scorecard based their self-assessments on a combination of evidence from evaluation reports, evaluation guidance documents, capacity building and awareness initiatives, as well as gender mainstreaming policies⁷.

As shown in Figure 3, 33 out of 42 reporting entities in 2020 (almost 80 per cent) employed the scorecard. 45 per cent of scorecard users “exceed requirements”, 42 per cent report “meeting requirements”, and only 12 per cent report “approaching requirements” (Figure 4). Overall, this represents a substantial improvement in ratings, as close to 88 per cent of scorecard users reached the established benchmark in 2020, compared with 76 per cent in 2019.

⁷ Entities that did not conduct any evaluations in 2020 (but have in previous years), used their 2019 rating, if available, as advised in the EPI technical note.

Figure 3. Disaggregated results for UN-SWAP reporting entities in 2020: by scorecard use and rating (N=42)

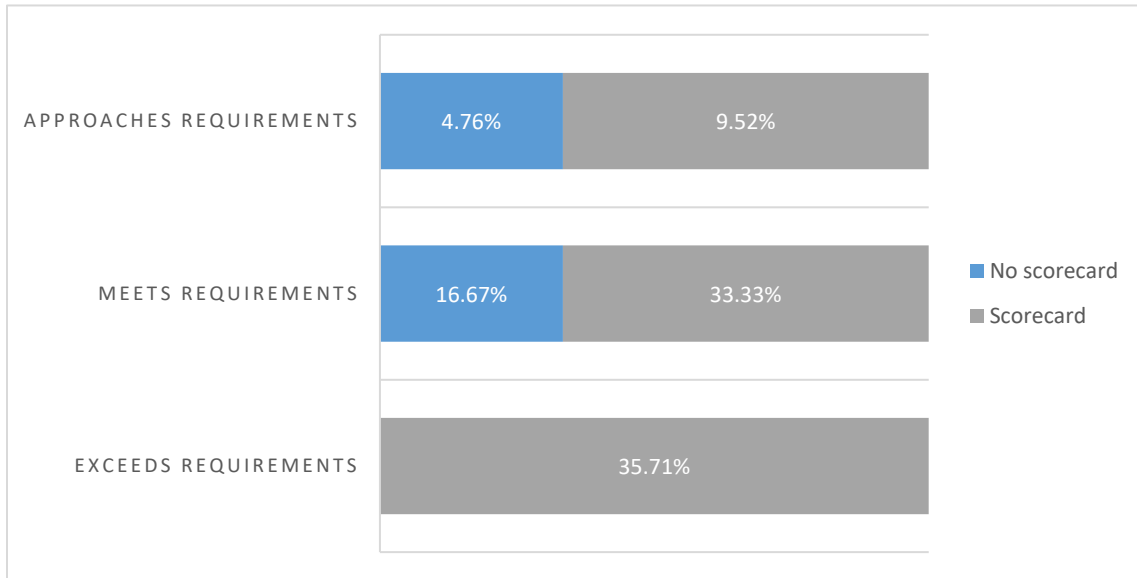
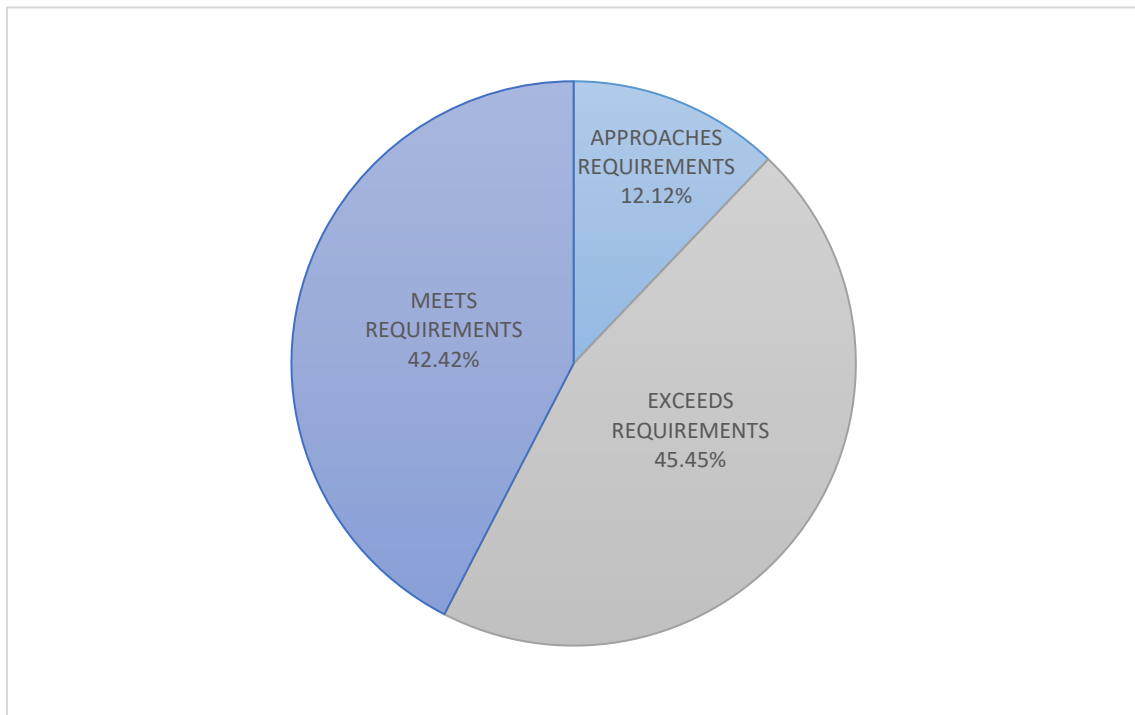


Figure 4. Disaggregated results for UNEG Scorecard users 2020: by rating (N=33)



Insight 3: There is an increase in the number of evaluations included in the EPI assessment sample and an overall improvement in the quality of reporting.

Table 3 presents an overview of the number of evaluations that were included by entities in the reporting/scoring sample. While most include five or fewer evaluations, in comparison to 2019, more entities employed a larger sample of 16 or more evaluations. The variation in sample sizes is not unexpected, as there is a wide diversity of mandates, capacities, number of evaluations conducted, and types of evaluation across entities.

Table 3. Disaggregated results for UN-SWAP EPI reporting entities in 2020: by number of evaluations in reporting sample (N=42)

EPI Rating	Number of evaluations					
	1-2	3-5	6-10	11-15	16-20	21 or more
Exceeds	ESCWA*, UNCDF*	ECE, ECLAC, OHCHR, WIPO			UNESCO, IFAD	FAO, IOM, UNFPA, UNICEF, UNDP, UN Women, WFP
Meets	DPPA, DMSPC*, DSS, ITC, OCHA, OLA, OSRSG- SVC*, UNOCT, UNRWA*	DESA, UNAIDS, UNCCD, UNCTAD, UN- HABITAT*, UNITAR*	ESCAP, IAEA OIOS	UNHCR, UNODC, WHO		
Approaches	ECA*, IMO	UNV*			UNIDO	ILO, UNEP

*Entities did not upload scorecards to the UN-SWAP portal

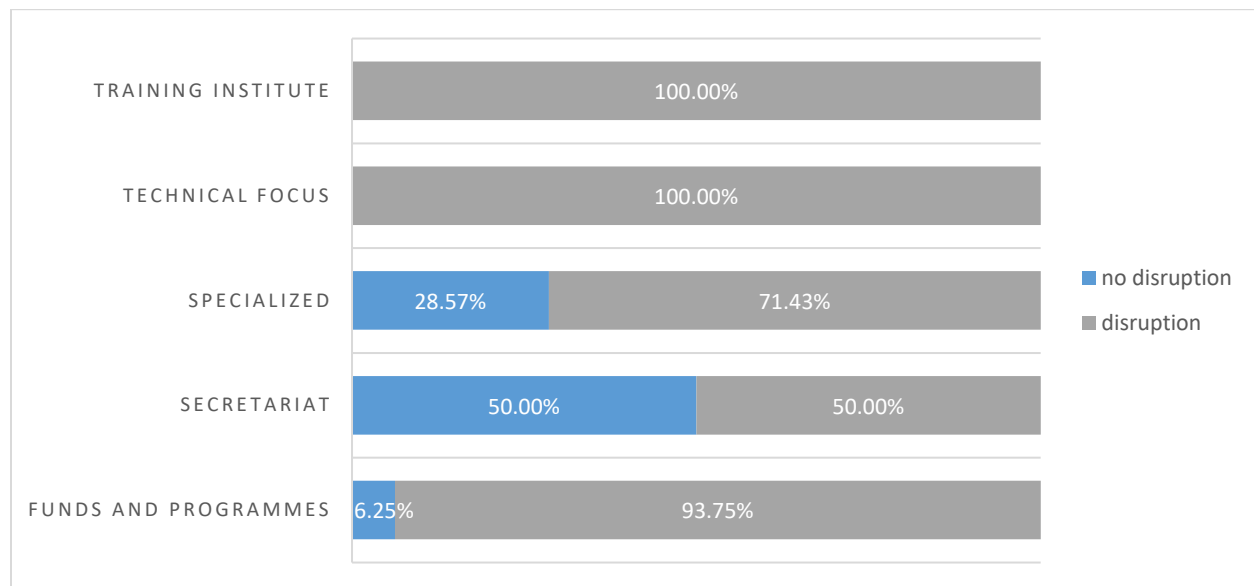
Note: The Global Environment Facility (GEF) is not an official UN-SWAP entity. However, as part of its continued commitment to gender-responsive evaluation, GEF voluntarily reports annually against the UN-SWAP EPI by strictly applying the UNEG-endorsed reporting process. As per the scorecard for 2020, GEF's aggregate rating is "approaches requirements" based on scores from nine evaluations.

Based on reporting feedback, the EPI technical note was updated in 2020 to clarify the use of the scorecard and requirements to assign additional points for commissioning a corporate evaluation of gender mainstreaming. In 2020, an overall improvement in the quality of reporting was seen, with fewer entities seeking clarifications on the criteria. In terms of review modality, nine entities employed external/independent assessments and two engaged in peer-reviews. Where possible and resource permitting, more entities should be encouraged to move from self-reporting on the EPI to other modes of assessment in the future.

Insight 4: COVID-19 related disruptions were indicated by nearly three fourths of reporting entities and are likely to continue to affect evaluation activities in 2021.

Close to three-fourths of all reporting entities indicated disruptions to their evaluation activities due to the COVID-19 pandemic. Figure 5 below presents the reported disruptions by entity type.

Figure 5. Disaggregated results for UN-SWAP reporting entities in 2020: by reported disruption to evaluation activities due to COVID-19 pandemic (N=42)



The most commonly cited disruptions were delays and postponement of evaluation activities. Owing to travel restrictions, most entities reported having to rely on remote data collection and somewhat restricted samples. This disproportionately affects gender responsive evaluation, as women often have unequal access to information technology and may be harder to reach using remote survey techniques.

The shift from face-to-face interviews to virtual formats also affects the length and nature of insights. Low participation rates in online surveys are another issue and can result in self-selection bias among respondents. Validation workshops were also reported as having been cancelled due to COVID-19, possibly affecting the extent to which findings were co-produced with implementers and beneficiaries. Entities also acknowledged the impact of the additional burden placed on staff owing to the pandemic and video conference fatigue. These factors are likely to continue to impact the quality of evaluations conducted in 2020 and 2021. Going forward, evaluations should clearly acknowledge these limitations and also propose ways to mitigate and address some of these barriers.

Good practices to advance integration of gender equality in evaluations

The quality of the EPI reporting in 2020 continued to be high, with detailed reports highlighting good practices and innovative efforts to support gender mainstreaming in evaluation. In addition, many good practices were highlighted related to adapting evaluation methods to the ongoing Covid-19 pandemic.

A. Several entities reported good practices in efforts to further integrate and mainstream gender equality in **quality assurance processes and evaluations**.

WFP included a chapter on **gender in the Research and Analysis Guide for Country Strategic Plan Evaluations** and additional guidance for Country Strategic Programme Evaluations has been drafted describing existing WFP frameworks and tools on gender and how they may be used in evaluations.

UNFPA's management kit for country programme evaluations includes guidance on integrating gender equality and women's empowerment in evaluation. The guidance will be rolled-out in 2021, presenting an opportunity to further strengthen the quality of evaluations in general and the integration of gender equality elements in particular. In 2020, the UNFPA Evaluation Office also updated the evaluation quality assurance and assessment system to include a specific criterion on the integration of gender equality in evaluation which directly mirrors the language of the evaluation performance indicator of the UN-SWAP 2.0.

UNDP Independent Evaluation Office also fully incorporated the **UN-SWAP EPI into its online QA system** in November 2020, which means that all decentralized evaluations that are quality assessed online hereafter are also assessed against the UN-SWAP EPI. UNDP also launched Evaluation Awards for excellent decentralized evaluations, including one award for gender-responsive evaluation.

In 2020, the **UNAIDS** Evaluation Office developed a **checklist to review the quality of UNAIDS evaluation** reports and this includes the UN-SWAP indicator.

For the first time in 2020, **UNITAR** applied the **UNDP-developed GRES Gender Results Effectiveness Scale** to an evaluation and foresees to continue doing so in the future. UNITAR will continue to share the **"Guidance on Integrating Gender Considerations into Monitoring and Evaluation of UNITAR Programming"**, update the lessons learned database and particularly the category with regards to gender equality, as well as track the recommendations with regards to gender/women empowerment

In 2020, **UNEG** members continued to engage a wide range of **internal and external stakeholders** in supporting and strengthening the gender responsiveness of evaluations by conducting and following up on evaluations on gender mainstreaming, conducting trainings and issuing new knowledge products.

The 2019 Evaluation and Audit Report of **WIPO's Policy on Gender Equality** was a major milestone and provides an assessment of the results of the implementation of the Policy between 2014 and 2019. The report identifies six recommendations, which are currently under implementation by WIPO.

UNESCO conducted a **corporate (global) evaluation to assess corporate performance on gender mainstreaming**. While all corporate evaluations had some degree of gender integration and analysis, the decentralized evaluations showed larger variation depending on the nature of the programme or project

being evaluated and the extent to which a gender dimension had already been incorporated as part of the design of the project.

The **FAO** gender team provided the office with a **Framework for Harmonizing Gender Analysis**, which identifies evaluation questions, indicators, and tools to be included in all evaluations (in line with the UNEG gender-related norms and standards).

UN Women finalized a working paper and guidance note on evaluating impact on gender equality and the empowerment of women, as well as a rapid assessment tool to assess gender equality and the empowerment of women results in humanitarian contexts and a good practices booklet for gender-responsive evaluations in the context of the Beijing Declaration and Platform for Action and the Sustainable Development Goals (SDGs).

FAO's Office of Evaluation carried out a **gender-stock taking exercise in 2020**, to assess the extent of integration of gender equality perspectives in the evaluations and delivered three webinars on gender in English, Spanish, and French.

The **IOM** evaluation office continued to implement **training initiatives** designed in 2018, one for internal evaluators training and the other one on a comprehensive approach of monitoring and evaluation, both including elements related to the promotion of gender coverage in evaluations.

At the global level, the **UNICEF** Evaluation Office will organise a **learning session** to discuss the lessons that emerged from the assessment of the 2019 portfolio and the office will continue supporting Regional Offices to organize training and information-sharing webinars with country offices on the areas of evaluation methods, lessons learned, and UN-SWAP principles. Furthermore, evaluation managers will continue to strengthen the requirements in the evaluation Terms of Reference for these elements and select evaluators who have proven experience at integrating evaluation principles, particularly SWAP.

OIOS plans to organize **specific training on gender responsive evaluations, conduct a self-evaluation** of the performance on gender mainstreaming and issue a gender inclusion checklist to be used in evaluation assignments.

WHO is currently implementing its first corporate evaluation of the integration of gender, equity and human rights in the work of WHO (to be completed in Quarter 3 2021).

B. Additional good practices related to adaptation of evaluation methods due to ongoing Covid-19 pandemic

The UNFPA Evaluation Office issued a **specific guidance and framework** to regional and country offices on how to adapt evaluations during the COVID-19 crisis. These principles emphasizes the need to pay particular attention to the most-at-risk groups and those furthest behind with a gender lens.

The UNICEF evaluation function issued a technical note outlining the implications of the COVID-19 crisis for ongoing evaluation work and clarifying the role the evaluation function should play in these challenging times. UNICEF also supported **real time assessments** and rapid reviews of measures being undertaken in 20 countries to respond to COVID-19.

UN Women IES produced a **pocket tool for managing evaluations** during the COVID-19 pandemic. Due to the increased use of virtual data collection, the tool highlighted upholding relevant ethical protocols, using gender-responsive methods for remote data collection, storing and protecting sensitive data and sharing findings in accessible and user-friendly ways. UN Women also undertook two rapid assessments to identify key lessons and provide evidence for UN Women’s COVID-19 response.

In order to ensure full representation of all stakeholders (women, men, boys, and girls), UNODC developed **COVID-19 evaluation guidance** for programme managers and evaluation teams. The guidance included methodologies to promote a no-one left behind and participatory approach throughout the evaluation process while using remote data collection.

UNDP published several guidelines to help programme units in implementing evaluations during COVID-19. Due to the overall impact of the crisis, many programme units have revised their evaluation plans and **updated evaluation methodology**, i.e., collecting data remotely and conducting virtual stakeholder interviews, in order to ensure a principle of “do no harm” and the safety of staff, consultants and national stakeholders.

OIOS led a **real-time evaluation** of the UNRWA’s operational response to COVID-19, developed a COVID-19 response evaluation protocol, and produced a synthesis of guidelines for UN evaluation under COVID-19, as a guide for UN entities to adjust their evaluations in the context of the pandemic.

The UNAIDS Evaluation Office **adjusted evaluation approaches** and methods to minimize the potential impact of evaluations on national stakeholders. For all evaluations that were initiated since the COVID-19 outbreak, an iterative approach was adopted to regularly identify and confirm the feasibility and risks of each subsequent stage of an evaluation. In practice, most activities were conducted remotely, using national consultants and virtual communication technologies.

UNCDF’s evaluation unit adapted its work to the COVID-19 crisis and response. Among the criteria for selecting evaluation teams, importance was given to teams making primary use of national consultants. All evaluation TORs were updated with the latest **safety guidance and emerging needs** due to COVID-19 have led to changes to objectives and evaluation questions.

3. Way forward

2020 was the third year of reporting on the revised UN-SWAP 2.0. While the quality of the EPI reports has improved overall, and an increasing number of entities are using the UNEG endorsed scorecard, more entities should commission independent, or peer-review based assessments in the future. Entities have also reported integrating UN-SWAP indicators into their internal quality assessment frameworks, which will serve to further enhance the quality of gender mainstreaming in evaluations.

The COVID-19 pandemic will continue to impact evaluations conducted in 2020 and 2021. While most agencies reported disruptions owing to the pandemic, many have adjusted evaluation methods, issued guidance documents, and suggested ways to undertake and improve the quality of remote evaluations. There is a good opportunity for entities to learn from each other’s experiences and use the newly developed resources on conducting remote evaluations.

As in previous years, evaluation units of various entities have produced knowledge products on different aspects of gender responsive evaluations, including guidance documents for decentralized evaluations, pocket tools, rapid assessments, and gender frameworks for evaluations.

In 2021, the UNEG working group on Gender, Disability and Human Rights plans to undertake a gap analysis of key guidance documents, including the UNEG guide *on Integrating Human Rights and Gender Equality in Evaluations* to identify potential improvements such as disability inclusion and the new evaluation criteria.