

# Independent Quality Assessment of UNODC Evaluation Reports 2018

## Synthesis Report

Prepared for:  
Independent Evaluation Section, UNODC  
by John Mathiason and Ann Sutherland  
March 2019

## CONTENTS

1. INTRODUCTION.....	3
2. METHODOLOGY .....	3
3. FINDINGS .....	4
3.1 EQA Ratings.....	4
3.2 General Strengths and Improvements of the Evaluation Reports.....	6
3.3 General Weaknesses of the Evaluation Reports .....	7
3.4 Mainstreaming Human Rights and Gender Considerations .....	8
3.5 Best Practices Observed in Evaluation Reports .....	11
3.6 Improvements from Draft EQAs .....	12
4. ASSESSING CONTRIBUTION TO SDGs.....	13
5. RECOMMENDATIONS .....	15
Annex 1. List of evaluations reviewed.....	18
Annex 1. UNODC EQA Template as Used in the Review .....	19

## 1. INTRODUCTION

The Independent Evaluation Section (IES) is leading and guiding evaluations in order to provide objective information on the performance of United Nations Office on Drugs and Crime (UNODC). Evaluation is one of the key factors in the Secretary-General's accountability reforms that are being considered by the General Assembly in 2019.<sup>1</sup> UNODC is one of the only Secretariat departments that has its own evaluation Section. As such it is a major indicator of progress.

As a member of the United Nations Evaluation Group (UNEG), the IES is following its Norms and Standards. IES's work is based on three pillars: 1) National Evaluation Capacity Building and SDGs; 2) Evaluation results; 3) Evaluation knowledge products, communication and innovation.

As part of its efforts to ensure the office's independent evaluations (pillar 2) are providing credible information to inform planning processes, the IES has commissioned independent evaluation quality assessments of evaluation reports produced since 2014. The IES also seeks to improve evaluation practice and to create better mechanisms for tracking and using the knowledge gained from evaluations as part of ongoing organizational learning.

Building on the previous Evaluation Quality Assessment (EQA) report, this document synthesizes the EQA results of all published UNODC evaluation reports in 2018 (the list of reports is provided in Annex 1) and makes comparisons with EQAs since 2014/-2017.

This assignment was carried out from mid-January 2019 to mid-March of 2019 by two independent consultants - Dr. John Mathiason (Team Leader) and Ann Sutherland (Team Member). Both have extensive experience in conducting evaluations and meta-evaluations for international organizations. They are the Managing Director and Principal Associate, respectively, for Associates for International Management Services (AIMS).

## 2. METHODOLOGY

**EQA Template:** The first deliverable on the annual evaluation quality assessment review, was to assess the most recent EQA template and to propose any needed changes. The reviewers compared UNODC's template with those that used for other organizations, such as UNFPA and UNICEF, and have proposed only three changes that are shown in red in the attached template. The first is to note whether an EQA was applied to the draft report. The second, responding to the recommendation last year that the assessment of methodology is more robust, is to ensure that the data collection methods can clearly show the causal connection between outputs and outcomes. And the third is to simplify the gender UN-SWAP<sup>2</sup> scorecard to conform with what has now been agreed by the UN system (UN-SWAP 2.0), by eliminating one of the sub-criteria. The revised EQA template used for the assessment is shown in Annex 2.

**2018 EQA Process:** Based on this, the reviewers then examined the quality of all of the evaluation reports published during 2018. The total number of evaluations for the year was 16. Four were in-depth evaluations and 12 were independent project evaluations.

---

<sup>1</sup> [Eighth progress report on the accountability system in the United Nations Secretariat: strengthening accountability under the new management paradigm \(A/73/688\)](#)

<sup>2</sup> <http://www.unwomen.org/-/media/headquarters/attachments/sections/how%20we%20work/unsystemcoordination/un-swap/un-swap-2-framework-and-technical-guidance-en.pdf?la=en&vs=1406>

To ensure consistency of the review process, one quarter of the reports (four) were selected for assessment by both reviewers, including the three in-depth evaluations. The reviewers then compared their comments and scores for each criterion as well as the overall score. In all cases, the overall scores were the same for both reviewers. There were minor differences in criteria scores and there were non-material differences in comments. The differences were discussed and resolved. The remaining reports were each rated by one person with assignments being based on each reviewer’s area of expertise and language skills.

**Consideration of SDGs and HRGE:** The reviewers were requested to pay specific attention to the extent to which the Sustainable Development Goals and issues of Human Rights and Gender Equity were considered as part of evaluation processes. This report includes a separate chapter on each of these areas.

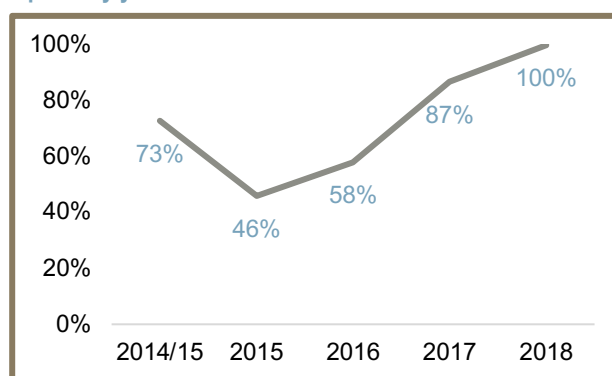
### 3. FINDINGS

#### 3.1 EQA Ratings

The 16 published evaluation reports in 2018 were rated highly and reveal the continuous improvement in the quality of UNODC evaluations since 2015<sup>3</sup>. Both Figure 1 and Table 1 show this trend.

The numerical scoring that was introduced to the EQA template for review of the 2017 reports, and again used in 2018, brings an additional level of preciseness to the ratings. The ratings are based on the overall scores of each report: Very Good (90%+), Good (70-89%), Fair (50-69%), and Unsatisfactory (<50%).

Figure 1: Percent of combined Very Good and Good reports by year



In 2018, five reports (31%) received a Very Good rating and no reports were rated as Fair or Unsatisfactory. The 2018 scores ranged from a low of 77% to a high of 96%. Particularly notable is that the average score was 86, which is excellent and 6 points higher than in 2017.

Table 1: Overall rating of 2018 reports compared with previous EQA cycles

	Unsatisfactory	Fair	Good	Very Good	Total
# of Reports - 2018	0	0	11	5	16
% of Reports - 2018	0%	0%	69%	31%	100
# of Reports - 2017	0	2	8	5	15
% of Reports - 2017	0%	13%	54%	33%	100%
# of Reports - 2016	0	8	7	4	19
% of Reports - 2016	0%	42%	37%	21%	100%
# of Reports - 2015 <sup>4</sup>	0	12	9	1	22

<sup>3</sup> It should be noted that since 2016 the IES includes the EQA template as an attachment to introductory emails to all evaluators. The EQA criteria and template are also found in the [Evaluation Handbook: Guidance for designing, conducting and using independent evaluation at UNODC](#) (October 2017).

<sup>4</sup> This included reports published from June through December 2015.

% of Reports – 2015	0%	53%	41%	5%	100%
# of Reports – 2014/15 <sup>5</sup>	0	9	22	2	33
% of Reports – 2014/15	0%	27%	67%	6%	100%

As with previous years, there were some notable differences by criterion. Table 2 shows that Recommendations tended to have lower ratings. In 2016, the lower ratings were found for the Executive Summary and Reliability sections. In 2017, Presentation & Structure rated the most highly. In 2018 Lessons Learned ranked highest.

**Table 2: Report rating by criteria**

	Unsatisfactory	Fair	Good	Very Good
Presentation/Structure	0	0	8	8
Executive Summary	0	1	13	2
Context & Purpose	0	2	7	7
Scope & Method	0	0	11	5
Reliability	0	0	8	8
Findings	0	1	8	7
Conclusions	0	2	11	3
Recommendations	0	6	9	1
Lessons Learned	0	1	3	12
Consideration of GEEW	0	2	7	7

Table 3 shows the scores by type of report: Independent Project Evaluations – midterm and final – and programmatic In-depth Evaluations. In 2018 the patterns were similar by type.

**Table 3: Rating by type of evaluation**

Type of evaluation	Good	Very Good	Grand Total
Independent Final	6	3	9
Independent Midterm	2	1	3
Indepth	3	1	4
<b>Grand Total</b>	<b>11</b>	<b>5</b>	<b>16</b>

Table 4 shows the scores by subject area. The highest percent of evaluations commissioned were in the area of Criminal Justice (44%). There were too few reports in the other areas to draw conclusive findings about which had the strongest reports, although the two (of two) undertaken for Crime Prevention and the one (of one) undertaken for Research were all rated as Very Good.

**Table 4: Rating by subject of project**

Subject	Good	Very Good	Grand Total
Anti-Corruption	1	0	1
Crime Prevention	0	2	2
Criminal Justice	5	2	7
Drug Trafficking	2	0	2

<sup>5</sup> This included the first batch of EQA assessments, which considered reports published from January 2014 through May 2015.

Organized Crime	3	0	3
Research	0	1	1
<b>Grand Total</b>	<b>11</b>	<b>5</b>	<b>16</b>

### 3.2 General Strengths and Improvements of the Evaluation Reports

As mentioned above, the evaluation reports for 2018 showed improvement from previous years with all receiving a rating of Good or Very Good. There were five primary strengths observed across those reviewed:

**Consistency with DAC and UNEG Norms:** As was the case for the past four years, the reports generally conformed to the accepted norms and guidelines. The DAC is in the process of reviewing and updating the criteria for assessing evaluation quality, but this should not substantially change how UNODC reports are reviewed. Several of the main changes being considered – such as more focus on gender equity and human rights, contribution to SDGs, partnerships, design rigour, and learning – are already incorporated into UNODC evaluation guidance and the EQA process. However, both should be updated once the DAC revisions are finalized.

**Inclusion of Gender Analysis:** This section was generally strong with 14 of the 16 reports rated as Very Good or Good. Further discussion is provided in the “Mainstreaming of Human Rights and Gender” section below.

**Lessons Learned:** This section of the reports received the highest score in 2018 with all but one report receiving a rating of Very Good (12) or Good (3). On the whole, the lessons learned sections showed well more general conclusions from the evaluation.

**Executive Summaries:** This section was highlighted as being the weakest section of the 2016 reports but has shown marked improvements over the last two years. For 2018, all but one of the 16 reports received a section score of Very Good or Good. The most significant improvement has been that these sections are more concise and less frequently exceed the maximum length of four pages. The most frequently seen issue with the 2018 Executive Summaries is omissions in stating the main purpose and target audience of the evaluations. In some cases, the object of the evaluation was not described. There is also a continuing trend of the emphasis, and most of the space, being given to summarizing the findings of each criteria rather than to the overall conclusions of the evaluation process. This is consistent with the DAC norms of showing results by criterion, however the absence of strong conclusions can make it more difficult for the reader to easily get a sense of the overall accomplishments, strengths and weaknesses of the evaluation object.

#### Other Areas of Noticeable Improvement:

- **Reliability of data:** This was another identified weakness in the 2016 evaluations with only seven of 17 reports having this section rated as Very Good or Good but where there has been continuous improvement. For 2018, half of the evaluations were Very Good, with the rest being Good.
- **Inclusion of visual aids:** Most reports included tables and figures to highlight key findings and analysis, although more could be done in this regard.
- **Systematic organization of Findings:** In most reports, the evaluation questions are stated for each criterium and a summary of the finding for each criterium is provided. This makes it very clear what the evaluation is addressing and what the main findings are.
- **HRGE:** There is generally more substantive discussion of the extent to which both human rights and gender equality issues were or were not addressed by the object of the

evaluation.

### 3.3 General Weaknesses of the Evaluation Reports

**Recommendations:** As noted above, this was the weakest section with six of the 16 reports receiving a Fair rating and only one receiving a Very Good rating. The most common issues were that (a) recommendations were not clustered or prioritized, (b) there were no recommendations related to SDGs, and (c) there was no indication that stakeholders were consulted as part of their development. A notable improvement from previous years is that most reports had recommendations that addressed any flaws in the project data systems.

There are areas in which UNODC evaluations could continue to improve upon that were also identified in previous EQA processes. These include:

**Challenges in making causal connections:** Given that a primary intent of most evaluations is to assess progress towards intended results, it is important for the evaluation process to show causal links between programme/project activities and the different levels of results. This is particularly the case for UNODC evaluations as these frequently are complex to conduct because of the topic and/or region. Many evaluators note that the initiative's logical framework is inadequate, and some then go on to suggest that it is not possible to assess effectiveness or impact, or they end up having overly descriptive rather than evidence-based findings. In three cases this year, the evaluators noted a logframe existed but did not include it in the report or annexes, which makes it difficult to determine what was expected to happen that would be measured by the evaluation. Only half of the evaluations in 2018 received a "yes" on Scope & Methodology question c. "*Methodology allows for drawing causal connections between output and expected outcomes*". This was a continuation of the problem found in previous years. In such cases, there should be some attempt to set up a logframe that is clear about what the evaluation process is measuring. At a minimum, the evaluators should articulate the proposed chain of results or program theory. This practice should be suggested in reviewing the inception reports.

**Methodological rigour:** There are two issues here. One is the use of relatively few data collection methods - most evaluations rely on document review, interviews, and field visits as data collection methods. When surveys are included, the response rate is typically too low for the results to be valid. Another concern is the lack of systematic data analysis (or description of how this was done). A notable improvement from previous years is that data is more frequently broken down by gender and by stakeholder group.

**Evidence of stakeholder consultation in developing conclusions and recommendations:** There are three EQA criteria that assess the extent of stakeholder participation in the evaluation process. Good practices call for stakeholder involvement in providing feedback on preliminary findings and in the formulation of conclusions and recommendations and UNODC ToRs are generally explicit about the need for engaging Core Learning Partners (CLPs) throughout the process. Although most all UNODC evaluations include consultation with a range of relevant stakeholders as part of data collection, few reports articulate how stakeholders may have been further involved. In fact, only six of the 16 evaluations received a "yes" on "*High degree of participation of internal and external stakeholders, including the Core Learning Partners, throughout the evaluation process is planned for and made explicit*" in the scope and methodology section. And, in terms of recommendations, only five evaluations received a "yes" on "*Reflect stakeholders' consultations whilst remaining balanced and impartial*" while four received a "no".

**Consideration of SDGs:** Fourteen of the 16 evaluations mentioned SDGs. However, as was the case in 2017, in most reports the connection and contribution of the initiative to the SDG is

described in very general terms, often just naming those goal(s) that are most relevant. Section 4 provides more analysis and guidance for how SDGs could potentially be addressed.

**Other common shortcomings of evaluation reports and processes:**

- Not providing information about evaluators and their suitability for the assignment;
- Not providing the dates/timeframe of the evaluation process;
- Use of interview protocols that do not appear to have been adapted for different stakeholder groups.

### 3.4 Mainstreaming Human Rights and Gender Considerations

As with previous years, all evaluation reports included dedicated sub-sections within the Findings for Human Rights and Gender. Table 5 shows that the average scores for considering GEEW have improved according to the UN-SWAP criteria. The SWAP tool assesses the extent to which gender equality and the empowerment of women (GEEW) is integrated into evaluation processes. In the past there were four criteria. In 2018, UNEG decided to reduce the number of GEEW criteria from four to three in order to simplify and clarify the scoring process. UNODC’s EQA template was adjusted for this change.

Each criterion is rated on a scale of 0 - 3 with 0 being awarded when there is no integration of gender, 1 when gender issues are partially integrated, 2 when it is satisfactorily integrated, and 3 when gender is fully integrated.

**Table 5: Average scores for the integration of GEEW (UN-SWAP)**

Quality Assessment Criteria	Average Score (0-3)			
	2015	2016	2017	2018
a. GEEW is integrated in the evaluation scope of analysis and indicators are designed in a way that ensures GEEW-related data will be collected.	1.125	1.63	1.93	2.44
b. Evaluation criteria and evaluation questions specifically address how GEEW has been integrated into design, planning, implementation of the intervention and the results achieved.	1.625	1.63	2.2	N/A
c. Gender-responsive evaluation methodology, methods and tools, and data analysis techniques are selected.	.875	1.37	1.87	2.25
d. Evaluation findings, conclusions and recommendations reflect a gender analysis.	1.875	1.84	2.2	2.31
Average Overall Score (out of 12)	5.5	6.53	8.2	N/A
<b>Average Overall Score (out of 9)</b>	N/A	N/A	N/A	<b>6.94</b>

Two of the four In-depth Evaluations (RAB and GLOY 09) both received GEEW scores that Meets Requirements. Notably, seven of the 12 Independent Project Evaluations received Meets Requirements scores – RER/F23, GLOT 63, GLOZ 82, KENZ 04, MMRZ 39, UZB/U57, and PSEX 02. The latter evaluation, which assessed Forensic Human Resource and Governance Development Assistance to the Palestinian Authority, was the only report to receive an overall score of nine (the highest possible).

The extent to which evaluation reports met each of the criteria is shown in Table 6. As can be seen, the reports were most successful in including GEEW as part of Criteria A. Scope & Indicators.



**Table 6: Number of reports achieving scoring requirements for each criterion**

	A. Scope/Indicator	B. Methodology	C. Findings/Conclusions/Recommend
Not at all integrated	0	0	0
Partially integrated	1	0	1
Satisfactorily integrated	7	12	9
Fully Integrated	8	4	6

Table 7 provides the overall EQA rating and the GEEW category scores for each report. It shows that the majority of 2018 reports (9 of 16) were rated as Good/Met Requirements both in their overall rating and in their GEEW rating. The average rounded overall GEEW score for the 2018 reports is 7.0 which shows that, overall, UNODC evaluations are Meeting Requirements which is the highest category. The average rounded score for all of the three criteria was 2. The table also includes the average scores (at the individual evaluation level and not rounded) for the respective criteria for the 2017 reports, and in each case the scores improved for the 2018 reports.

**Table 7: GEEW scores by evaluation report**

Project Number	Overall EQA Rating	Scores for each criterium			GEEW Total Score
		A	B	C	
XAPA 10	Very Good	2	2	2	6
RER/F23	Good	2	3	2	7
Cluster XAC/Z60 etc.	Good	2	2	3	6
GLOY09	Good	3	3	2	8
Cluster RAB	Very Good	3	2	3	8
XLAY08	Good	1	2	2	5
XAW/Z28	Good	2	2	2	6
GLOT63	Good	3	2	2	7
GLOZ82	Good	3	2	3	8
INDA03	Good	2	2	2	6
KENZ04	Very Good	3	3	2	8
MEXZ56	Good	2	2	1	5
MMRZ39	Very Good	3	2	3	8
GLOR35	Good	2	2	2	6
PSEX02	Good	3	3	3	9
UZB/U57	Very Good	3	2	3	8
<b>Average score 2018</b>		2.44	2.25	2.31	6.94
<b>Avg rounded score 2018</b>		<b>2</b>	<b>2</b>	<b>2</b>	<b>7</b>
Average score 2017		1.93	1.87	2.2	N/A

Observations about the way that each criterium was addressed in the 2018 evaluations is as follows:

A. Scope and Indicators: All reports included gender as part of the scope of the evaluation, and over half identified the process as being gender responsive. This suggests that the importance of including a gender lens is increasing being emphasized in preliminary discussions with evaluation consultant commissioned by UNODC. In 2018, all evaluations included criteria and questions pertaining to gender. However, it was not always possible to see if gender-related indicators were specified, given that logical frameworks or evaluation matrixes were not consistently included in the reports.

B. Methodology: As in 2017, the strongest reports were those that described how gender was to be analyzed and how the methodology was designed to be gender responsive. In several instances the methodology included a phrase or sentence about a gender responsive methodology being used but there was no explanation about how the evaluation was actually carried out in a gender responsive manner. In all cases, the stakeholders consulted as part of the evaluation were gender disaggregated but not all reports included this figure in the main report; the information could only be found in the annexed list of stakeholders consulted. There were no cases where the evaluators explicitly explained how differences in power relations amongst evaluation participants were addressed, even in instances where focus groups included both men and women (women were always in the minority). However, the RER/F23 report did note that women were approached separately during site visits for their perspectives on the workplace environment.

C. Findings, Conclusions and Recommendations: All reports included gender-related findings, usually about the extent to which gender was integrated into the design and implementation of the evaluation object. These sections were usually brief, but some provided a very thorough analysis (for example, GLOY 89, GLOR 35, PSEX 02, Cluster RAB, MMRZ 39). Some considered gender within multiple criteria (for example, PSEX 02, Cluster RAB, INDA 03, MMRZ 39). Most all had gender-related conclusions and recommendations.

**Good GEEW Practices Observed:** The reviewers observed the following factors as contributing to higher GEEW scores and being demonstrative of good practice over the past two years:

- Having a dedicated gender expert as part of the evaluation team;
- Being explicit in the Methodology section about how the evaluation process was gender responsive;
- Conducting a HRGE (human rights and gender equality) evaluability assessment during the inception phase of the evaluation;
- Including an assessment of gender in multiple criteria under Findings, not just in HRGE;
- Extending the analysis of GEEW in the evaluation beyond an assessment of the extent to which gender was integrated into the design and implementation of the intervention to also consider how gender could have been better incorporated into the design and how this would benefit the project/programme.

One of the reports that was particularly strong in analysing gender issues was the **RAB Cluster** evaluation. Notable were that an HRGE evaluability assessment was conducted during the inception phase of the evaluation and a dedicated gender expert was included in the evaluation team.

The **RER/F23** report was unique in that the Methodology section highlighted the persistence of the evaluation team in pursuing gender questions in the focus group sessions even when male participants seemed less interested in this part of the agenda. Furthermore, it was noted that as gender disaggregated statistics of training participants were not kept by the project, the evaluators did a retrospective analysis by reviewing the names of attendance records.

The **GLOY 09** report was also noteworthy. While GEEW was not a feature of the programme's design, the evaluators made a good effort to collect data to examine current and potential use of GEEW. The evaluation scope included looking at ways "to strengthen HRGE in the GP" and the

collection of this data was prioritized. The team included an HRGE focal point, but the methodology also stated that all team members were "researching, analysing and writing about HRGE issues throughout the evaluation process" so that "responsibilities for HRGE were mainstreamed into the evaluation process".

The evaluators of the **MMRZ 39** project constructed a theory of change for the intervention, highlighting its role in reducing gender inequalities. The inception phase included a stakeholder analysis to ensure a diverse range were included in the study, and approximately 40% of people eventually consulted were women. Sampling criteria for field visits took into account villages where the 'presence of women is stronger'. Other strengths included the Background section of the report addressing how gender was integrated into the project, and that a gender analysis was included under five criteria - Design, Relevance, Partnership, Effectiveness, HRGE. Gender was also addressed in multiple places within Conclusions and Recommendations, and was highlighted as part of Lessons Learned.

**PSEX 02** is a project that has sexual and gender-based violence as a major focus, and its evaluation was exemplary for being gender responsive. Gender was fully addressed as part of the scope, both in respect to GBV and broader aspects of gender equity. The methodology was stated as being gender-sensitive and inclusive, and was explicit in how this was accomplished. For example, a mixed methods approach was used, including storytelling, and there was reference to how GBV data was handled to ensure anonymity and to the need to collect data in ways that were sensitive to the socio-cultural context. A variety of stakeholders were consulted and data collection tools enabled gender disaggregation. Gender analysis was included under several criteria, and there was substantial context provided as part of the HRGE criteria. The conclusions were gender focused and several recommendations addressed furthering the GBV objectives of the project.

**Composition of evaluation team:** For the 2018 EQA review, data were analyzed according to the composition of the evaluation teams. Most included both men and women, but there was a curious result when only women were on the team. In those four cases, the average GEEW score was lower than for either evaluations with mixed teams or only men. It may be the case that some evaluators routinely use inclusive practices throughout the evaluation process, and that this has become so normalized that they do not see it necessary to highlight how their work aligns with gender responsive principles.

### 3.5 Best Practices Observed in Evaluation Reports

**Showing causal connections** - XAPA 10 *Strengthening and Enhancing the Capacity of Law Enforcement Officials in Combating Child Sex Offenders in Cambodia, Lao PDR and Viet Nam*. In this case, the evaluator was diligent in showing the relationship between project outputs and results. The Findings for both Efficiency and Effectiveness were organized by outcome area, and the Effectiveness section was further structured by target indicator. A Table of Progress and Outputs of the Project annexed to the report provided additional detail connecting activities to results.

**Robust methodology and refined logical framework** - GLOU34, GLOX64, GLOV44, GLOV20 and MEXX35 *Global research projects of the Research and Trend Analysis Branch* (RAB Cluster). This evaluation is notable for its robust design that also enabled a clear assessment of causal connections. It is one of the few evaluations for which the logical framework was revised to improve its evaluability and was adapted to also function as an evaluation matrix. The team used multiple methods to ensure validation of findings including an extensive set of interviews and a well-implemented survey of disaggregated users of knowledge products, training workshops and technical assistance. Throughout the Findings section, the various methods that confirmed the finding are listed. As noted earlier, the evaluation is also exemplary for its in depth of analysis of HRG issues.

The importance of having a strong logical framework for assessing what is supposed to have happened and causal connections cannot be overstated. Four of the 2018 evaluation reports did not include a logframe (or portion thereof), even though in all cases it was noted these existed for the respective interventions. Several reports included an evaluation matrix; this is good practice however it is not a variable assessed in the EQA template.

**Good use of visual aids and formatting** – The *RAB Cluster* evaluation also makes exemplary use of visual aids throughout Findings section to highlight key data and data sources. Additionally, the report helpfully provides hyperlinks in the main text to relevant annexes.

**Careful Use of Data and Analysis:** *UZB/U57 Umbrella Project in Support to the Implementation of ROCA Program*. This is a small-scale innovative project that was designed to provide a supplement to the regular work of the Regional Office for Central Asia (ROCA). It was rated very good because of the care by which the evaluator obtained data and analyzed it in order to answer the questions posed in the terms of reference.

**Description of stakeholder involvement** - *XAW/Z28 Support to the ECOWAS Regional Action Plan on illicit drug trafficking, related organized crime and drug abuse in West Africa*. This is notable for succinctly, yet clearly, stating the role of the 12 members of the CLP.

**Analysis of Human Rights** - *INDA 03 Strengthening Anti-Corruption Institutions in Indonesia* has a strong analysis of the link between human rights and anti-corruption. *GLOR 35 EU-Nigeria-UNODC-CTED Partnership Project II: Assisting Nigeria to strengthen rule of law-based criminal justice responses to terrorism* is also exemplary for its human rights analysis.

### 3.6 Improvements from Draft EQAs

The previous EQA process included the review of four draft evaluation reports. The final versions of three were reviewed as part of the 2018 EQA process with the KGZ/K50 report being reviewed with the 2017 batch. As can be seen in Table 8, in all cases the overall rating improved. The sections most frequently highlighted as needing improvement in the draft versions were Context & Purpose, Scope & Methodology and Recommendations. In the final reports, the overall ratings for these sections improved by one or two levels, with the numerical scores increasing by 7 to 25 points.

**Table 8: Differences between draft and final report EQAs**

	Draft Report EQA rating	Final Report EQA rating	Main issues from draft reports
GLOZ 82 Doha Declaration	Fair (score 64)	Good (score 89)	Context/Purpose and Scope/Methodology sections weak (Unsatisfactory). SDGs not mentioned. The final report addressed these issues and the report was only 1 point below a Very Good rating.
KGZ/K50 Counter Narcotics Service	Good (score 79)	Very Good (score 90)	Sections for improvement were Executive Summary, Scope/Methodology and Recommendations (all rated as Fair). For the final report, these sections were rated respectively as Good, Very Good and Very Good.
GLOY 09 Paris Pact	Fair (score 69)	Good (score 83)	Context/Purpose, Scope/Methodology and Lessons Learned sections weak (Unsatisfactory). SDGs not addressed. For the final report, all previously Unsatisfactory sections received a Good rating but SDGs were still not addressed.
RAB Trend & Analysis Branch	Good (score 86)	Very Good (score 93)	Sections for improvement were Context/Purpose and Recommendations (all rated as Fair). In the final report both of these sections were rated as Good.

## 4. ASSESSING CONTRIBUTION TO SDGs

One of the issues that evaluations began to address and will have to address further in 2019 is the role of evaluation in the review process of the Sustainable Development Goals, especially Goal 16, to promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels. The review process is organized around Voluntary National Reviews (VNRs) that are considered by the High-Level Political Forum (HLPF) of the Economic and Social Council and, once every fourth year, to the General Assembly. Goal 16 will be considered at the 2019 Session of the HLPF.

As currently designed, the reviews are to be based on indicators, whose determination takes place via the UN Statistical Commission. There are three levels of indicators, called Tier I, II and III. Tier I indicators are those for which there is a consensus and reasonable assurance that they could be measured if the data were collected by the country. Tier II Indicators are conceptually clear, with established methodology and standards available but data are not regularly produced by countries. Tier III Indicators are those for which there are no established methodology and standards or methodology/standards are being developed/tested. Six of the goals and 44 indicators are specifically a concern of UNODC. UNODC is a custodian (or co-custodian) of 13 indicators, meaning that it is responsible for developing, maintaining and encouraging data collection on the indicators. Ten of these are Tier II or Tier III indicating that more work needs to be done to make them useable for the review process.

One problem that is evident from a review of the indicators is that for most of those for which UNODC is the principal, especially in Goal 16, there are no real indicators of targets that have been agreed that measure how national institutions function. The exceptions are 3.3.1 on HIV incidence, 3.8.1 on coverage of health services which would include substance abuse treatment, 16.1 on homicides and 16.4 on money-laundering. For the others, there are few indicators for which data are currently being collected by national statistical systems, or where the indicators really measure whether the target has been achieved. As a result, reporting on these has been limited.

This can be seen in the first stage of the review process that has been the presentation of voluntary national reviews (VNRs) at the High-Level Segments in 2016, 2017 and 2018. One hundred fifty-four countries presented VNRs. While there are annual appraisals of the VNRs, they do not provide target-based data and they varied in content and approach. The analysis of which of the UNODC indicators the VNRs considered in 2017, shown in Table 8 and presented in the 2017 review of UNODC evaluations, shows significant variation and confirms the finding that many indicators are not really available. The target with the most references in the VNRs was 5.2 on violence against women (35 percent), largely because there are five organizations that collect this data systematically in many countries. The next largest number of references is to 3.3.1 on HIV/AIDS (30 percent) because WHO collects this information in many countries. Only two other indicators had references greater than 20 percent, and those are 3.5.1 dealing with health services (a Tier III on which UNODC is working with WHO) and 16.1 on homicides (a Tier I indicator).

As can be seen from Table 9, the analysis from 2017 also shows that the number of references to Goal 16 is very limited.

**Table 9. Goal 16 Indicators mentioned in VNRs**

Indicator	Whether Mentioned		
	No	Yes	Total
16.1 - Violence	32	11	43
16.3 - Prosecution	35	8	43

16.4 - Illicit flows	37	6	43
16.5 - Corruption	35	8	43
16.6 - Transparent government	36	7	43
16.7 - Participation in Decision-making	39	4	43
16.8 - Transparent public services	40	3	43
16.10 - Human rights violations	36	7	43
16.a - Independent Human Rights institutions	39	4	43
16.b - Harassment	40	3	43

The table suggests that a special effort will be needed to ensure that States report on progress, including by assisting them to collect data, both by determining what to measure and when that is clear, how to collect the data.

In this context, evaluation, which is part of the SDGs in paragraph 74 (g) and (h) of General Assembly resolution 70/1 that adopted the SDGs, should be a major element in reporting, especially where Tier I indicators do not exist. Because evaluations collect data, especially at the country-level, about what has happened in subject areas, and why, they can be an important tool in SDG reporting. To test this, SDG targets that were covered by UNODC evaluations in 2018 were examined as shown in Table 10. All but two of the evaluations made specific reference to the SDGs. As in past years, most (10) were connected with Goal 16, with the largest connected with Target 16.4 on *significantly reduce illicit financial and arms flows, strengthen the recovery and return of stolen assets and combat all forms of organized crime*. The second, with eight evaluations, was 16.3 to *Promote the rule of law at the national and international levels and ensure equal access to justice for all*. The number of targets is much larger than in 2017, but one factor is that a single evaluation, MMRZ39 (Sustainable Livelihoods and Development in Myanmar 2014-2019) referred to 17 targets.

**Table 10. SDG Targets Covered by UNODC Evaluations, by Type, 2018**

SDG Targets	In-Depth Evaluations	Independent Project Evaluations	Grand Total
1.4 Equal rights to economic resources		1	1
2.1 Access to safe, nutritious and sufficient food		1	1
2.3 Double the agricultural productivity and incomes of small-scale food producers		1	1
2.4 Ensure sustainable food production systems		1	1
3.5 Strengthen the prevention and treatment of substance abuse	2	1	3
4.7 All learners acquire the knowledge and skills needed to promote sustainable development		1	1
5.2 Eliminate all forms of violence against all women and girls	2	2	4
5.a Undertake reforms to give women equal rights to economic resources		1	1
8.2 Achieve higher levels of economic productivity		1	1
8.3 Promote development-oriented policies that support productive activities		1	1
8.5 Achieve full and productive employment and decent work for all women and men		1	1
8.7 Eradicate forced labour, end modern slavery and human trafficking	1	0	1
10.1 Achieve and sustain income growth of the bottom 40 per cent of the population		1	1
10.2 Empower and promote the social, economic and political inclusion of all,		1	1
11.7 Provide universal access to safe, inclusive and accessible,		1	1

green and public spaces			
15.1 Ensure the sustainable use of terrestrial and inland freshwater ecosystem		1	1
15.2 Promote the implementation of sustainable management of all types of forests		1	1
15.3 Combat desertification,		1	1
15.4 Ensure the conservation of mountain ecosystems		1	1
15.7 Take urgent action to end poaching	1		1
15.9 Integrate ecosystem and biodiversity values into national and local planning		1	1
16.1 Significantly reduce all forms of violence	2		2
16.2 End abuse, exploitation, trafficking and all forms of violence against and torture of children	2	2	4
16.3 Promote the rule of law at the national and international levels and ensure equal access to justice for all	2	3	5
16.4 Significantly reduce illicit financial and arms flows and combat all forms of organized crime	3	3	6
16.5 Substantially reduce corruption and bribery	2	2	4
16.6 Develop effective, accountable and transparent institutions	1	2	3
16.7 Ensure responsive, inclusive, participatory and representative decision-making		0	0
16.a Strengthen relevant national institutions to prevent violence and combat terrorism and crime	2	2	4
16.b Promote and enforce non-discriminatory laws and policies		1	1
<b>Grand Total</b>	<b>20</b>	<b>15</b>	<b>55</b>

This suggests that, if organized and applied, UNODC evaluations, like those of other organizations, could be applied in the SDG review process. How this could be done, however, is something that is still being discussed at the UN system level. Partly it is because the main instrument for review, the UNDAF, does not have a consistent evaluation process. However, one could be developed if there was an agreement among UN system agencies as part of the reform of country-level activities and if this is endorsed by governments. This is currently underway in the Economic and Social Council. In practice, in countries where it has projects, UNODC is on the country teams and as such has a role in the UNDAFs. Evaluations have been done in 19 countries since 2013 as well as 14 evaluations of regional and sub-regional projects, and eleven global projects. Forty-six of these were focused on Goal 16 subjects. In that sense, UNODC can contribute to UNDAF evaluations where its projects exist.

One essential element is to ensure that an appropriate reference to the relevant SDG and target for which the evaluation is relevant (and, in fact, projects need to specify this in their design). Another is that the IES should make an effort to capture and catalogue data on SDG results in its databases as a matter of routine.

## 5. RECOMMENDATIONS

The main recommendations for the IES to consider are:

- 1. Ensure that cause-effect links are clearly reflected in the evaluations.** The heart of an effective RBM evaluation is to show that the outputs produced by the programme or project have a clear connection with the outcomes achieved. For this to happen, there has to be an adequate logical framework described so that methods of collecting data can be effective for assessing attribution. A theory of change (or results chain) is also important for depicting the interconnections between outcomes and the overall objective.

At the institutional level, IES should continue to advocate for all UNODC programmes and projects to have clear and up-to-date logframes and theories of change. It is also recommended that the adequacy of both, but particularly the logframes, be revised if necessary during the evaluation inception phase by the evaluators in collaboration with IEU. These tools then need to consistently be included as part of the evaluation reports.

- 2. Continue to place more emphasis on robust methodologies:** It is important that the methodology used adequately facilitates answers to the evaluation questions. Evaluators should be encouraged to include more quantitative data collection and analysis processes. If surveys are not feasible, interview protocols and other tools should be designed to elicit at least some responses that can be more easily quantified through content analysis processes. Evaluators also need to be more explicit about how they have analyzed the data collected through both quantitative and qualitative processes. Additionally, evaluators should be encouraged to routinely disaggregate responses by stakeholder groups as well as by gender - it is unlikely that all stakeholders hold the same views and the varying perspectives should be illuminated in the Findings.
- 3. Encourage more explicit descriptions of gender-responsive practices:** As part of the inception stage, it is suggested that IES review the relevant sections of the SWAP scoring system with each evaluation team. Evaluators should also be reminded up front of the need to clearly articulate in the methodology section of the report the steps they have taken to ensure the evaluation process used was gender-responsive.
- 4. Encourage more explicit descriptions of stakeholder participation:** Good practice calls for stakeholder involvement in evaluation processes and this should be built into the evaluation design. Evaluators should also be reminded about the need to be explicit about how the Core Learning Partners (and other stakeholders as relevant) have been involved including their contribution to the development of the report's conclusions and recommendations.
- 5. Addressing other ongoing issues that affect EQA scores:**
  - Visual aids: Although evaluators are increasingly including tables and graphs to present information, use of data visualization techniques is still minimal. More could be done in this regard.
  - Information about evaluators and their suitability for the assignment: This information is frequently not provided.
  - Dates/timeframe of the evaluation process: This information is frequently not provided.
- 6. Improve assessment of SDGs:** UNODC should continue to take a lead in showing how evaluation can improve the assessment of the SDGs. This can involve working to improve indicators that are currently Tier II or III, or where the current indicators do not measure what the target should achieve. It should also work in country-teams where UNODC has projects, to ensure that evaluations are consistent and that SDGs for which UNODC is responsible are included.
- 7. Further refinement of the EQA template:** Two revisions are proposed in order to reduce some redundancy in the template.
  - There are currently three sub-criteria that assess stakeholder involvement – in *Scope & Methodology*, *Conclusions*, and *Recommendations*. These could be reduced to a single sub-criterion, likely best placed under *Structure, Completeness and Clarity*, that would assess whether there was full participation of key stakeholders (normally the CLP) in the critical points of evaluation design, conclusions and recommendations.



- There is duplication in assessing whether the limitations of the data and data collection are identified and addressed (currently included under *Scope & Methodology* and under *Reliability of Data*). It is suggested this only remain under the latter criterion.
- 8. Institute EQAs of Draft Reports:** Based on the extent of improvement seen in the EQA scores of the draft and final reports (for those select cases where draft reports were assessed), consideration should be given to instituting EQAs of the final draft of evaluation reports on a more systematic basis.

## Annex 1. List of evaluations reviewed

All reports are published on UNODC website at [www.unodc.org/unodc/en/evaluation/reports.html](http://www.unodc.org/unodc/en/evaluation/reports.html)

Project Number	Project Title
XAPA 10	Strengthening and Enhancing the Capacity of Law Enforcement Officials in Combating Child Sex Offenders in Cambodia, Lao PDR and Viet Nam
RER/F23	Drug law enforcement systems for criminal intelligence collection, analysis and exchange
XAC/Z60, TAJ/E24, TAJ/H03, RER/H22, XAC/K22	In-Depth Thematic Cluster Evaluation of the projects: XAC/Z60, TAJ/E24, TAJ/H03, RER/H22, XAC/K22
GLOY09	Paris Pact Initiative Phase IV- A Partnership to Combat Illicit Traffic in Opiates Originating in Afghanistan
GLOU34, GLOX64, GLOV44, GLOV20 and MEXX35	Global research projects of the Research and Trend Analysis Branch
XLAY08	Proyecto de Apoyo a la Reducción de la Demanda de Drogas Ilícitas en la Comunidad Andina - PREDEM
XAW/Z28	Support to the ECOWAS Regional Action Plan on illicit drug trafficking, related organized crime and drug abuse in West Africa
GLOT63	Support to Crime Prevention and Criminal Justice Reform
GLOZ82	Global Programme for the Implementation of the Doha Declaration: towards the promotion of a culture of lawfulness
INDA03	Strengthening Anti-Corruption Institutions in Indonesia
KENZ04	Evaluation of the Police Reform Programme in Kenya
MEXZ56	Fortalecimiento de las estructuras de seguridad y justicia en el Estado de Coahuila de Zaragoza
MMRZ39	Sustainable Livelihoods and Development in Myanmar 2014-2019
GLOR35	EU-Nigeria-UNODC-CTED Partnership Project II: Assisting Nigeria to strengthen rule of law-based criminal justice responses to terrorism
PSEX02	Forensic Human Resource and Governance Development Assistance to the Palestinian Authority
UZB/U57	Umbrella project in support to the implementation of ROCA Program

## Annex 1. UNODC EQA Template as Used in the Review

## UNODC Evaluation Quality Assessment (EQA) Template

General Project Information		
Project/Programme Number and Name		
Thematic Area		
Geographic Area (Region, Country)		
Relevant SDG(s)		
Approved project/programme budget of the time of the evaluation (USD)		
Type of Evaluation (In-Depth/Independent Project; Final/Midterm; Other)		
Cost of Evaluation (USD)		
Evaluation Team (# of men/# of women; names of team members)		
Date of Evaluation (from MM/YYYY to MM/YYYY)		
Date of Evaluation Report (MM/YYYY)		
Quality Assessment conducted on/by		
EQA provided for draft report (Y/N; if Y, indicate rating)		

<b>OVERALL QUALITY RATING:</b>	
<b>SUMMARY:</b>	

Quality Assessment Criteria	Yes No Partial	Assessment Levels: Very Good - Good - Fair - Unsatisfactory Meets Criteria: Y = Yes N = No P = Partially N/A- Not Applicable	
<b>1. Structure, Completeness And Clarity Of Report</b>		<b>RATING:</b>	<b>#DIV/0!</b>
a. Format (headings, font) accords to IEU Guidelines and Templates for Evaluation Reports.	<Select one>		
b. Structure accords to IEU Guidelines for Evaluation Reports with the following logical sequence: List of acronyms; Executive Summary; Summary Matrix of Findings, Evidence and Recommendations; Introduction (Background and Context, Evaluation Scope and Methodology, Limitations to the Evaluation); Findings (Relevance, Efficiency, Partnership and Cooperation, Effectiveness, Impact, Sustainability, Human Rights and Gender Equality/Mainstreaming, as well as Design and Innovation if in ToR); Conclusions; Recommendations; Lessons Learned.	<Select one>		

c. Objectives stated in the terms of reference are adequately addressed.	<Select one>	
d. Issues of human rights and gender equality/mainstreaming are adequately addressed.	<Select one>	
e. Report is easy to read and understand (i.e. written in an accessible non-technical language appropriate for the intended audience).	<Select one>	
f. Language is empowering and inclusive avoiding gender, heterosexual, age, cultural and religious bias, among others.	<Select one>	
g. Report is free from grammar, spelling, or punctuation errors.	<Select one>	
h. Visual aids, such as maps and graphs, are used to convey key information.	<Select one>	
i. Composition of Evaluation Team is included and has gender and geographic expertise. Preferably it is gender balanced and includes professionals from countries or regions concerned.	<Select one>	
j. Annexes include at a minimum: evaluation terms of reference; logic model and/or evaluation matrix; list of persons interviewed and sites visited; list of documents consulted; evaluation tools used.	<Select one>	
<b>2. Executive Summary</b>		<b>RATING:</b> #DIV/0!
a. Written as a stand-alone section that provides an overview of the evaluation and presents its main results.	<Select one>	
b. Generally follows the structure of: i) Purpose, including intended audience(s); ii) Objectives and brief description of intervention; iii) Methodology); iv) Main Conclusions; v) Recommendations.	<Select one>	
c. Summary Matrix presents only the key and most important recommendations from evaluation report.	<Select one>	
d. Findings, sources and recommendations in the Summary Matrix are clear and cohesive, and specify the stakeholder to whom they are addressed.	<Select one>	
e. Maximum length 4 pages, excluding the Summary Matrix.	<Select one>	
<b>3. Evaluation Context And Purpose</b>		<b>RATING:</b> #DIV/0!
a. Clear description of the project evaluated is presented.	<Select one>	
b. Logic model and/or the expected results chain, and /or program theory (that at a minimum identifies and links objectives, outcomes and indicators of the project) is clearly described.	<Select one>	
c. Connection with Sustainable Development Goals is clear.	<Select one>	

d. Context of key cultural, gender related, social, political, economic, demographic, and institutional factors are described, and the key stakeholders involved in the project implementation and their roles are identified.	<Select one>	
e. Project status is described including its phase of implementation and any significant changes (e.g. to strategies, logical frameworks) that have occurred.	<Select one>	
f. Purpose of evaluation is clearly defined, including why it was needed at that point in time, what information is needed, and the target audience.	<Select one>	
<b>4. Scope And Methodology</b>		<b>RATING:</b> #DIV/0!
a. Evaluation scope is clearly explained including the main evaluation criteria, questions and justification of what the evaluation did and did not cover.	<Select one>	
b. Transparent description presented of methodology applied, including how it was designed to address the evaluation purpose, objectives, questions and criteria.	<Select one>	
c. Methodology allows for drawing causal connections between output and expected outcomes.	<Select one>	
d. Methods are appropriate for analysing gender equality/mainstreaming and human rights issues identified in evaluation scope; methodology takes into account power relations during an evaluation process; is inclusive and participatory.	<Select one>	
e. Data collection methods and analysis, and data sources are carefully described, as are the rationale for selecting them. <b>Processes are adequate for measuring outcomes set out in logic model.</b> The limitations are also clearly described. Reference indicators and benchmarks are included where relevant.	<Select one>	
f. Sampling frame clearly described and includes area and population to be represented, rationale for selection, mechanics of selection including whether random, numbers selected out of potential subjects, and limitations of sample.	<Select one>	
g. High degree of participation of internal and external stakeholders, including the Core Learning Partners, throughout the evaluation process is planned for and made explicit. When there are thematic or approach gaps (i.e. gender equality/mainstreaming) among stakeholders, external key informants were invited for consultation.	<Select one>	
<b>5. Reliability of Data</b> <i>To ensure quality of data and robust data collection processes</i>		<b>RATING:</b> #DIV/0!

a. Triangulation principles (using multiple sources of data and methods) were applied to validate findings.	<Select one>	
b. Qualitative and quantitative data sources were used, and included the range of stakeholder groups and additional key informants (when necessary) defined in evaluation scope.	<Select one>	
c. Limitations that emerged in primary and secondary data sources and collection processes (bias, data gaps, etc.) are identified and, if relevant, actions taken to minimize such issues are explained.	<Select one>	
d. Evidence provided of how data was collected with a sensitivity to issues of discrimination and other ethical considerations.	<Select one>	
e. Adequate disaggregation of data by relevant stakeholder undertaken (gender, ethnicity, age, under-represented groups, etc.). If this has not been possible, it is explained.	<Select one>	
<b>6. FINDINGS AND ANALYSIS</b> <i>To ensure sound analysis and credible findings</i>		<b>RATING:</b> #DIV/0!
<i>Findings</i>		-
a. Are clearly formulated and presented	<Select one>	
b. Are based on rigorous analysis of the data collected; take into account any identified benchmarks.	<Select one>	
c. Address all evaluation criteria and questions raised in the ToR including relevance, efficiency, effectiveness, impact and sustainability, as well as UNODC's additional criteria of design, partnership and cooperation, innovation, and the cross-cutting themes of human rights and gender.	<Select one>	
d. Address any limitations or gaps in the evidence and discuss any impacts on responding to evaluation questions raised in ToR.	<Select one>	
e. Include findings related to the SDGs covered by the evaluation.	<Select one>	
f. Discuss any variances between planned and actual results of the project (in terms of objectives, outcomes, outputs).	<Select one>	
<i>Analysis</i>		-
a. Interpretations are based on carefully described assumptions.	<Select one>	
b. Contextual factors are identified (including reasons for accomplishments and failures, and continuing constraints).	<Select one>	

c. Cause and effect links between an intervention and its end results (including unintended results) are explained.	<Select one>		
d. Includes substantive analysis of human rights issues.			
<b>7. CONCLUSIONS</b>		<b>RATING:</b>	#DIV/0!
a. Take into consideration all evaluation criteria and questions, including human rights and gender equality/mainstreaming criteria.	<Select one>		
b. Have been formulated clearly, are based on findings and substantiated by evidence collected and go beyond the findings and provide a thorough understanding of the underlying issues of the project and add value to the findings.	<Select one>		
c. Convey the evaluators' unbiased judgement of the intervention.	<Select one>		
d. Developed with the involvement of relevant stakeholders.	<Select one>		
e. Present a comprehensive picture of both the strengths and weaknesses of the project.	<Select one>		
<b>8. RECOMMENDATIONS</b>		<b>RATING:</b>	#DIV/0!
a. Are clearly formulated, based on the conclusions, and substantiated by evidence collected.	<Select one>		
b. Address flaws, if any, in project's data acquisition processes.	<Select one>		
c. Are specific, realistic, indicate a time, are actionable, and of a manageable number.	<Select one>		
d. Are clustered and prioritized.	<Select one>		
e. Include recommendations related to the SDGs covered by the evaluation.	<Select one>		
f. Reflect stakeholders' consultations whilst remaining balanced and impartial	<Select one>		
g. Clearly identify who is responsible for action.	<Select one>		
<b>9. LESSONS LEARNED</b>		<b>RATING:</b>	#DIV/0!
a. Are correctly identified, innovative and add value to common knowledge.	<Select one>		
b. Are based on specific evidence and analysis drawn from the evaluation.	<Select one>		
c. Have wider applicability and relevance to the specific subject and context.	<Select one>		
<b>10. ASSESSMENT OF THE INTEGRATION OF GENDER EQUALITY AND EMPOWERMENT OF WOMEN (GEEW) for UN-SWAP</b>		<b>RATING:</b>	#DIV/0!
a. GEEW is integrated in the evaluation scope of analysis and indicators are designed in a way that ensures GEEW-related data will be collected.	<Select one>		



<del>b. Evaluation criteria and evaluation questions specifically address how GEEW has been integrated into design, planning, implementation of the intervention and the results achieved.</del>	<Select one>	
c. Gender-responsive evaluation methodology, methods and tools, and data analysis techniques are selected.	<Select one>	
d. Evaluation findings, conclusions and recommendations reflect a gender analysis.	<Select one>	
<b>Overall Score</b>	<b>#DIV/0!</b>	
<b>Overall Rating</b>	<b>#DIV/0!</b>	

## SCORING

Assessment Levels: Very Good (90%+) Good (70-89%) Fair (50-69%) Unsatisfactory (<50%)

Element Of The Evaluation	Points Per Category	Average score	Weighted score
Presentation And Completeness	10	#DIV/0!	#DIV/0!
Executive Summary	5	#DIV/0!	#DIV/0!
Evaluation Context And Purpose	5	#DIV/0!	#DIV/0!
Evaluation Scope And Methodology	10	#DIV/0!	#DIV/0!
Reliability Of Data	5	#DIV/0!	#DIV/0!
Findings And Analysis	30	#DIV/0!	#DIV/0!
Conclusions	10	#DIV/0!	#DIV/0!
Recommendations	15	#DIV/0!	#DIV/0!
Lessons Learned	5	#DIV/0!	#DIV/0!
Gender Equality and Empowerment of Women	5	#DIV/0!	#DIV/0!
<b>Total Maximum Score</b>	<b>100</b>		<b>#DIV/0!</b>
<b>GEEW scoring: 0 - 3 points = Misses requirements / 4 - 6 points = Approaches requirements / 7 - 9 points = Meets requirements</b>			