



UNEG
United Nations Evaluation Group

Meta-synthesis of United Nations Development Assistance Framework (UNDAF) Evaluations with a Gender Lens

December 2019



Commissioned by the United Nations Evaluation Group
Human Rights and Gender Equality Working Group



Acknowledgement

This synthesis was commissioned by the United Nations Evaluation Group (UNEG) Working Group on Gender Equality and Human Rights. Charlemagne Sophia Gómez worked as a consultant and we would like to express deep appreciation of her hard work and dedication towards producing a quality report.

Messay Tassew (United Nations Entity for Gender Equality and the Empowerment of Women – UN Women – and co-chair of the UNEG Working Group on Gender Equality and Human Rights), Michael Francis (United Nations Office of Internal Oversight Services – OIOS), Heather Bryant (United Nations Development Programme – UNDP), and Sabas Monroy (Office of the United Nations High Commissioner for Human Rights – OHCHR) and co-chair of the UNEG Working Group on Gender Equality and Human Rights) served as key players in this exercise and provided constant and helpful direction, comments and useful suggestions to improve both the methodological approach and content of the report.

Special thanks also to all those who participated in the production of this report, particularly Natalie Raaber (United Nations Population Fund – UNFPA), Ekaterina Sediakina-Rivière (United Nations Educational, Scientific and Cultural Organization – UNESCO), Marianne Schmitt (International Trade Centre – ITC), Katherine Aston (United Nations Office on Drugs and Crime – UNODC), Tami Aritomi (United Nations Children's Fund – UNICEF), Pietro Tornese (United Nations Capital Development Fund – UNDCF), Emilie Wiinblad Mathez (United Nations High Commissioner for Refugees – UNHCR), and all members of UNEG Working Group on Gender Equality and Human Rights for their invaluable comments and suggestions and time taken to review the report.

We are also grateful to all the Resident Coordinators and members of the Development Coordination Office (DCO) who took the time to contribute to this undertaking, both via Skype and email, with their important insights and valuable comments. We would also like to express our appreciation to all the evaluators who were consulted for their responses and for sharing their experiences and recommendations to improve the evaluation process in the future.

Foreword

The 2030 Agenda for Sustainable Development established at its core the concepts of “leaving no one behind” and “reaching the furthest behind first.” In this respect, the United Nations Sustainable Development Cooperation Framework (hereafter “Cooperation Framework”) reinforces the mandate of the United Nations Development System (UNDS) to comprehensively and adequately address human rights and gender equality concerns. This should be done through, inter alia, clearly defined, gender-specific outcomes and outputs along with robust monitoring and evaluation.

The United Nations Evaluation Group (UNEG) is a frontrunner in advancing international evaluation practice and has developed several key guidance documents on how to ensure the mainstreaming of gender and human rights into evaluation. This meta-synthesis of evaluations of United Nations Development Assistance Frameworks (UNDAFs) – the predecessor of Cooperation Frameworks – offers a stark reminder of the need to improve the systems and processes for their management and quality assurance.

The insights and recommendations presented in this review represent a key opportunity to strengthen system-wide attention to gender equality and women’s empowerment in the core work of the United Nations Country Teams (UNCTs) and improve the overall quality and integration of gender equality dimensions in the new generation of Cooperation Framework evaluations.

I would like to take this opportunity to thank the UNEG Human Rights and Gender Equality Working Group for leading this very important exercise from its inception to the final outcome.

Susanne Frueh

Chair, United Nations Evaluation Group

A message from the UN Women Independent Evaluation and Audit Services

The timing of the release of this report will be significant as 2020 approaches. The year 2020 will mark the twenty-fifth anniversary of the adoption of the Beijing Declaration and Platform for Action and the twentieth anniversary of United Nations Security Council Resolution 1325, on women, peace and security, as well as herald a five-year milestone for achieving the Sustainable Development Goals (SDGs) of the 2030 Agenda for Sustainable Development. These anniversaries provide an opportunity to galvanize momentum to achieve gender equality and women's empowerment and increase the importance of gender-responsive evaluations as drivers to accelerate progress.

The synthesis indicates that while there have been significant efforts and progress, the pace of change is inadequate to attain an optimal level of quality and gender responsiveness of UNDAF evaluations – the predecessor of the United Nations Sustainable Development Cooperation Framework (UNSDCF).

Systematically addressing the bottlenecks will depend on intensified efforts of various actors. These include the capacity and ability of United Nations Country Teams (UNCTs) and Resident Coordinator Offices (RCOs) to design and monitor human rights and gender-responsive results matrices and relevant indicators; collection of disaggregated data and subsequent monitoring; enhanced system-wide capacity and coordination for Cooperation Framework evaluation processes and products with a gender lens; and adequate resource investment to engage gender-balanced and qualified evaluators.

I hope the insights presented in this synthesis constitute a relevant evidence base to reflect on the opportunities and challenges in improving the future generation of gender-responsive Cooperation Framework evaluations.

Lisa Sutton

Director, Independent Evaluation and Audit Services

UN Women

Acronyms & Abbreviations

CIS	→	Commonwealth of Independent States
ECA	→	Europe and Central Asia
EPI	→	Evaluation Performance Indicator
ERC	→	Evaluation Resource Center
GDI	→	Gender Development Index
GE	→	Gender Equality
GEWE	→	Gender Equality and Women's Empowerment
GRES	→	Gender Results Effectiveness Scale
HDI	→	Human Development Index
HR	→	Human Rights
LIC	→	Low Income Country
LGBTQ+	→	Lesbian, Gay, Bisexual, Transgender and Questioning Plus
LMIC	→	Low Middle Income Country
LNOB	→	Leave No One Behind
MDGs	→	Millennium Development Goals
RCOs	→	Resident Coordinators Officer
SDGs	→	Sustainable Development goals
TOR	→	Terms of Reference
UMIC	→	Upper Middle Income Country
UNDAF	→	United Nations Development Assistance Framework
UNSCDF	→	United Nations Sustainable Cooperation Development Framework
UNCTs	→	United Nations Country Teams
UNDP	→	United Nations Development Programme
UNDCO	→	United Nations Development Coordination Office
UN Women	→	United Nations Entity for Gender Equality and the Empowerment of Women
UNEG	→	United Nations Evaluation Group
UNDS	→	United Nations Development System
UN-SWAP	→	United Nations System Wide Action Plan

Glossary

Feminist Approach to Evaluation:

Def: Emphasizes participatory, empowerment and social justice agendas. It is based on feminist research, which in turn is based on feminist theory. Its main focus is gender inequalities that lead to social injustice. It seeks to ensure that women's experiences in evaluations are valued equally to those of men and does not treat women as a homogeneous group and recognizes others' gender identities. In addition, it encourages evaluators to use their empirical findings to advocate for social change.

Gender Blind:

Def: Results paid no attention to gender, failed to acknowledge the different needs of men, women, girls and boys or marginalized populations. (It is one of the five criteria of the Gender Results Effectiveness Scale).

Gender Development Index:

Def: The Gender Development Index (GDI) is the proportion of the Human Development Index (HDI) calculated separately for women and men using the same methodology as the HDI. It is a direct measure of the gender gap shown by the female HDI as a percentage of the male HDI.

Gender Inequality Index:

Def: The Gender Inequality Index (GII) is an inequality index that measures gender inequalities in three important aspects of human development: reproductive health, measured by the maternal mortality rate and adolescent birth rates; empowerment, measured by the proportion of parliamentary seats occupied by women and the proportion of adult women and men over 25 years of age with at least some secondary education; and the economic status, expressed as participation in the labour market and measured by the participation rate in the labour force of the populations of men and women aged 15 years or older. The GI is based

on the same framework as the Inequality-adjusted Human Development Index (IHDI), to better explain the differences in the distribution of achievements between women and men. The greater the value of GI, the greater the disparities between women and men and the loss of human development.

Gender Negative:

Def: Results had a negative outcome that aggravated or reinforced existing gender inequalities and norms. (It is one of the five criteria of the Gender Results Effectiveness Scale).

Gender Neutral:

Def: An outcome statement that neutralizes any reference to gender, and thus cannot discriminate on the basis of gender. (It is one of the four criteria of the Gender Ranking).

Gender Responsive:

Def: Results addressed differential needs of men or women and address equitable distribution of benefits, resources, status and rights but did not address root causes of inequalities in their lives. (It is one of the five criteria of the Gender Results Effectiveness Scale).

Gender Sensitive:

Def: Women, girls and/or gender issues/ needs are not the main focus of the outcome statement, but are mentioned in the outcome statement, footnotes, or explanation in the narrative to contribute to gender equality and women's empowerment. An outcome with a focus on human rights can also be considered gender sensitive (gender responsive). (It is one of the four criteria of the Gender Ranking).

Gender Specific:

Def: Women, girls and/or gender-related issues/needs are the focus of the outcome statement to promote gender equality and women's empowerment (gender responsive). (It is one of the four criteria of the Gender Ranking).

Gender Targeted:

Def: Results focused on the numeric value of equity (50/50) of women, men or marginalized populations that were targeted. (It is one of the five criteria of the Gender Results Effectiveness Scale).

Gender Transformative:

Def: Results contribute to changes in norms, cultural values, power structures and the roots of gender inequalities and discriminations. (It is one of the five criteria of the Gender Results Effectiveness Scale).

Gender@ Work Framework:

Def: The Gender@ Work framework, developed by an international collaborative, that helps organizations to build cultures of equality and justice, identifies the areas of change (individual change, informal change, formal change and systemic change) and the types of change (consciousness and awareness; access to resources and opportunities; policies, laws and arrangements; internal culture and exclusionary practices) that must occur to achieve transformative change.

Gender Results Effectiveness Scale:

Def: The Gender Results Effectiveness Scale (GRES) was created by the UNDP Independent Evaluation Office evaluation team to categorize the level of effectiveness or the quality of gender results, that is, it allows the results to be classified according to the type of gender change advanced and its effectiveness in transforming gender relations. The scale presents five categories of gender outcomes (gender transformative, gender responsive, gender targeted, gender blind and gender negative).

Applied Def: It has been used to determine the general approach of each report and to measure the exchange rate expressed in relation to the gender and human rights results captured in the main elements of the evaluation

Human Development Index:

Def: The Human Development Index (HDI) is a summary measure of the average achievement in the key dimensions of human development: a long and healthy life (assessed by life expectancy at birth), education (assessed by the average years of schooling for those over 25 years old plus the expected years of schooling for boys and girls of school age) and a decent standard of living (assessed by gross national income per capita). The HDI is the geometric mean of the normalized indexes for each of the three dimensions.

Gender Ranking:

Def: A gender measurement system that allows the results to be classified according to four areas (specific gender, sensitive gender, neutral gender and blind gender).

UN-SWAP EPI:

Def: The evaluation performance indicator (EPI) is one of the performance indicators developed as part of the accountability framework of the United Nations System-wide Action Plan (UN-SWAP) to implement the main Chief Executives Board for Coordination Policy (CEB/2006/2) on Gender Equality and the Empowerment of Women. It facilitates the evaluation of the degree to which the reports comply with the UNEG Standards and Gender-related Standards and demonstrate the effective use of the UNEG Guidance on integrating human rights and gender equality throughout all phases of the evaluation.

UNCT-SWAP Gender Equality Scorecard

Def: An accountability framework at the Country Team level for the United Nations System-wide Action Plan on Gender Equality and the Empowerment of Women.

Objective and audience

The meta-synthesis had two objectives: (i) to assess the gender and human rights responsiveness of a selection of United Nations Development Assistance Framework (UNDAF) evaluations; and (ii) to analyse key findings and recommendations of UNDAF evaluations for gender equality results.

This exercise comes at a time in which the United Nations Development System (UNDS) is undergoing ongoing reform, including to the way in which United Nations Country Teams (UNCTs) plan, programme and assess their work. Gender equality and women's empowerment (GEWE) has featured prominently in the primary objectives of UNCT work across the globe, but there has been limited evidence on the extent to which UNDAF evaluations to date are conducted in a gender-responsive manner and whether they assess the progress of human rights and gender-specific results. To address this gap, this meta-synthesis of UNDAF evaluations through a human rights and gender equality lens was commissioned by the United Nations Evaluation Group (UNEG) Working Group on Gender Equality and Human Rights.

Potential users of this report include the Development Coordination Office (DCO), Resident Coordinators' Offices (RCOs), the UNCTs and United Nations Regional Evaluation Groups, as well as donor agencies, civil society organizations and other national partners. It is envisioned that the report will contribute to the conduct of a new generation of more gender-responsive Cooperation Frameworks and United Nations Sustainable Development Cooperation Framework (UNSDCF) evaluations.

Methodology

This report answers a number of questions and divides them into two components aligning with the two key objectives of the meta-synthesis: i) questions relating to the analysis of the evaluation reports against the United Nations System-wide Action Plan (UN-SWAP) EPI; and ii) questions relating to the analysis and synthesis of the individual reports and

how they report on human rights (HR) and gender equality (GE)-related results and recommendations.

In order to respond to these two components, the meta-synthesis undertook a mixed-method approach, utilizing qualitative and quantitative data. The meta-synthesis largely relied on a comprehensive desk review of a sample of 50 evaluation reports conducted between 2015 and 2019. These evaluations covered the roll-outs of UNDAFs between 2010 and 2016. Accordingly, the design of the majority of the UNDAFs under assessment were concluded prior to the introduction of the Sustainable Development Goals (SDGs), and therefore the evaluation reports provided very little insight on alignment to the SDGs and the leave no one behind principle. The desk review of these evaluation reports was complemented by a significant literature review of articles and reports on gender and human rights evaluations. A small number of Skype interviews also took place with a sample of the evaluators (four) who had written the reports under assessment and interviews with six regional country offices in Africa, Asia and the Europe and Central Asia regions.¹

A survey was distributed to approximately 20 evaluators, to which eight persons responded.

In order to analyse the data, a number of analytical frameworks were utilized. The United Nations System-wide Action Plan (UN-SWAP) on Gender Equality and the Empowerment of Women evaluation performance indicator (EPI) criteria was used to determine the extent to which the evaluation reports had integrated HR and GE in the scope and design of the evaluation, its methodology and the findings, conclusions and recommendations.

To determine the quality and content of the GE results, the synthesis used two analytical frameworks: the Gender Results Effectiveness Scale (GRES) and Gender@Work Framework. Given the time and resource limitations, the synthesis focused on four key areas: governance, economic empowerment, environment/climate change and poverty.

¹ The Arab and Latin America regions were approached, but due to conflicting time schedules, interviews did not take place.

Executive Summary

Findings and conclusions

The meta-synthesis concluded that there is no single variable which effectively contributes to a HR and GE-responsive evaluation. While there are a number of factors which need to be improved upon in order to achieve more positive results, the quality of evaluators is essential in order to ensure a HR and GE-responsive evaluation. The desk review of the evaluation reports as well as the key informant interviews (KII) revealed that human resources plays an important part in the evaluation process, and limited human resources has often meant that UNDAF results matrices are poorly and hastily designed, thus making it more difficult for evaluators to extract the appropriate information and measure the impact at both outcome and output level. This is accentuated by the fact that 48 out of 50 reports cited weak monitoring and evaluation of the UNDAF results which rarely utilized HR and GE-sensitive indicators to measure the results. To this end, the design stage of both the results matrix and the Monitoring and Evaluation Framework are critical – without improvements in this area, reporting on progress of HR and GE results will be stalled. Furthermore, the average length afforded to the evaluators to conduct an evaluation tended to be around 25 days, leaving very little time for evaluators to visit the field and consult vulnerable groups and gather sufficient information based on evidence in order to draw judgments on the standard DAC/OECD evaluation criteria.

Upon examination of the UNDAF frameworks, it was concluded that there is an important correlation with the extent to which HR and GE is integrated in the evaluations and the extent to which the UNDAF is gender responsive. The meta-synthesis found that the UNDAFs which were classified as highly gender responsive (i.e. over 60 percent of their outcomes detailed in the results framework matrix were deemed gender responsive), tended to be ranked higher on the UN-SWAP EPI and thus were considered to integrate HR and GE to a more substantial degree.

The assessment showed that only a minority of UNDAF evaluations were gender responsive.

Four out of 50 evaluation reports (8 percent) met the requirements of the UN-SWAP EPI. A further 27 reports (54 percent) were found to have approached requirements. The integration of HR and GE into the design and the methodology of the evaluations had a propensity to be weaker than the reporting of programmatic results and findings. Approximately 19 out of these 27 reports could have perhaps met the requirements if the design and methodology had been more in line with HR and GE. The remaining 19 reports (40 percent) missed requirements.

The quality of the evaluation report correlates with the extent that HR and GE perspectives were integrated into the evaluation. The meta-synthesis revealed that nearly a quarter (12) of the evaluation reports in the sample were of poor quality. All 12 reports did not meet all the requirements. The reports which tended to be deemed higher in quality were more likely to integrate HR and GE. The meta-synthesis found that 67 percent of the reports which achieved a quality rating of moderately satisfactory or above either approached or met UN-SWAP EPI requirements. Quality assurance is therefore paramount to ensure that evaluations are up to the required standard and evaluations integrate HR and GE.

The failure of the evaluators to include HR and GE-responsive questions in the evaluation may have resulted from a weak terms of reference (ToR) as evaluators tended to mirror the scope of the ToR. In total, 60 percent of the ToR did not include a HR and/or gender scope. However, 54 percent did include at least one question on HR and/or GE. Fourteen out of 19 of these approached and/or met requirements, with five of them not meeting some requirements.

Only 40 percent of the evaluation reports either satisfactorily or completely integrated a HR and GE scope into the questions. In total, 19 (38 percent) reports included a HR or GE scope in the evaluation. However, 34 (68 percent) of the reports did include at least one type of question on HR and GE in the results matrix.

The second criterion on the UN-SWAP EPI which examines the extent to which the gender-responsive methodology, tools, methods and data analysis techniques are selected, was also found to be weak, with 30 reports (60 percent) either not integrating any aspect of HR and/or GE at all or only partially integrating them. The meta-synthesis found that the majority (94 percent) of the reports included a mixed methodology approach which is customary in HR and GE-responsive evaluations. Nonetheless, only 11 (22 percent) of the reports included disaggregated data, with much of this data being cited in the background section of the report. It is assumed that the reluctance of including disaggregated data is in part due to the difficulties encountered by both the Government and the UNCTs in collecting relevant data. HR and GE-sensitive indicators were also very rare.

The third criterion of the UN-SWAP EPI, which examines the extent to which the findings, conclusions and recommendations integrated HR and GE perspectives, was found to be the strongest of the three criteria.

Under the second component, utilizing the Gender Results Effectiveness Scale (GRES), it was concluded that the UNDAFs were deemed to contribute to the promotion of HR and GE, although the real impact is often difficult to gauge. The assessment of the four focus areas (governance, economic empowerment, environment/climate change and poverty) revealed that there were references to HR and GE in the programmatic results, and 77 percent of these were assessed as gender responsive, and thus examined the needs of women, girls and other vulnerable groups. A further 67 (25 percent) were deemed gender targeted and 32 (7 percent) gender blind. Only 2 percent of the references were found to be gender transformative. Environment/climate change had the highest number of gender-blind results.

Results which integrated HR and GE tended to be reported under governance, which accounted for the largest proportion of results out of the four focus areas under assessment. The different areas tended to use different spheres/levels of change to obtain their results. Governance and climate change would use the change in policies and laws to drive the majority of their changes, whereby economic empowerment and poverty focused the majority of their efforts on access to resources and opportunities. All four areas paid little attention to internal culture and exclusionary practices in order to harness a transformative change.

While many of the results were found to be gender responsive, transformative changes were only seen under nine results. Notwithstanding, many of these results had the potential to lead to transformative change, although this may require more time and a more strategic approach to ensuring adequate gender-responsive monitoring and reporting of the results.

There was little mention of vulnerable groups and women still feature very predominantly in what is deemed as a vulnerable group. The examination of the extent to which the UNDAF had been able to address the leave no one behind principle revealed that of the 443 references to HR and/or GE, only 189 made reference to at least one vulnerable/marginalized group – the majority of these (144) were found under governance. Nonetheless, upon further examination, the analysis revealed that the inclusion and integration of vulnerable groups were predominantly comprised of women (35 percent), such as rural and/or poor women or women and girls who were victims of domestic violence. A total of 14 percent of the references were made to youth and 29 percent for other vulnerable groups such as displaced people, refugees, the LGBTQ+ community and prisoners.

Recommendations

The key recommendations from the report are as follows:

- 1. Efforts should be made to ensure that human rights (HR) and gender equality (GE) are sufficiently mainstreamed across Cooperation Frameworks. This includes the need to guarantee more HR and GE-sensitive indicators in order to measure the impact the interventions have had. Efforts should concentrate on ensuring a balanced Cooperation Frameworks which allows HR and GE to be mainstreamed across all strategic priorities instead of a few key areas.**
- 2. UNCT staff capacity should be built and support provided to ensure expertise in HR and GE-sensitive design and monitoring of Cooperation Frameworks.**
- 3. The monitoring of the results should be carried out as stipulated in the UNSDCF guidelines while ensuring that HR and GE indicators are mainstreamed throughout the results. Cooperation Frameworks need to disaggregate beyond gender and ensure that different groups, such as persons with disabilities and the LGBTQ+ community, among others, are included.**
- 4. The evaluation report should be quality assessed through a gender lens utilizing the UN-SWAP EPI criteria. Furthermore, it is essential that all future Cooperation Framework evaluations are routinely uploaded in a publicly accessible repository.**
- 5. ToR should ensure a HR and GE-responsive evaluation by ensuring compliance with UNEG guidelines on integrating HR and GE in evaluations.**
- 6. RCOs/UNCTs should provide evaluators with support throughout the entire evaluation process to ensure that HR and GE are adequately integrated.**
- 7. UNEG could create a roster of evaluators with expertise in HR and GE-responsive evaluations.**
- 8. UNEG should update guidelines by producing a small FAQ on Cooperation Framework evaluations and a PowerPoint tutorial with best practices and sample gender equality and human rights questions.**
- 9. UNCTs should develop a gender strategy utilizing the implementation of the UNCT-SWAP Gender Equality Scorecard.**

Contents



- Acknowledgement.....I
- Acronyms & Abbreviations.....IV
- Glossary.....V
- Executive Summary.....VIII
- Contents.....XI
- Figures.....XIII
- 1. Introduction: Purpose and background.....1
 - 1.1 Background.....1
 - 1.1.1 Intended Use and users of the meta-synthesis.....2
 - 1.1.2 Context.....3
 - 1.2 Purpose and objectives of the Meta-synthesis.....3
 - 1.2.1 Approach and methodology.....4
 - 1.2.2 Sampling.....5
 - 1.2.3 Approach to meta-synthesis.....6
 - 1.3 Limitations.....7
 - 1.4 Structure of the report.....8
- 2. Key findings - the gender responsiveness of UNDAF evaluations.....9
 - 2.1 The impact of human rights/gender specific UNDAF’s on the integration of HR and GE in the evaluation reports.....9
 - 2.2 The UNCT’S commitments to gender and human rights.....11
- 3. Key findings - UN-SWAP EPI.....14
 - 3.1 Assessment of the gender responsiveness of the fifty UNDAF evaluations.....14
 - 3.2 Methodology.....14
 - 3.2.1 Integration of HR and GE perspective.....15
 - Criterion One - Scope and Design.....17
 - 3.3 The impact of the terms of reference on UNDAF evaluations.....18

3.5. The importance of ensuring evaluators with HR and GE experience.....	19
Criterion Two – Methodology.....	21
Criterion Three – Findings, Conclusions and Recommendations.....	24
IV: Key findings.....	27
4.1 Introduction.....	27
4.2. Overall assessment of the UNDAF’S contribution to HR and GE results.....	27
4.2.1. The contribution of the UNDAF’S on HR and GE.....	28
4.2.2. Key drivers of change on HR and GE.....	30
4.2.3. HR/GE Results under the four focus areas.....	34
4.3. The extent to which UNDAF’S have been able to address the leave no one behind principle.....	39
4.4. Extent to which collaborations and inter-agency cooperation including joint programmes and joint gender programmes contribute to greater efficiency on gender and human rights results.....	41
4.5. Evaluation Report Recommendations.....	42
4.6. Key challenges to ensuring a gender responsive evaluation.....	49
4.7. Twin Track Approach.....	52
V: Conclusions and recommendations.....	53
VI: Annexes.....	62
ANNEX I – Overall Assessment of the Fifty Evaluations under the UN-SWAP EPI.....	62
ANNEX II – Bibliography.....	65

Figures

Figure 1	→	UN SWAP EPI Criteria	1
Figure 2	→	Total number of UNDAF rollouts included in the Meta-Synthesis	4
Figure 3	→	Sample of UNDAF evaluations by geographic location	5
Figure 4	→	Universal sample	5
Figure 5	→	Overview of the reports in the universe	5
Figure 6	→	Methodological approach to Meta-Synthesis	7
Figure 7	→	Outcome Categorization Definition	9
Figure 8	→	Percentage of UNDAFs deemed to be gender specific and gender sensitive	10
Figure 9	→	Percentage of gender responsive priority areas of the UNDAFs under evaluation	10
Figure 10	→	Percentage of gender responsive outcomes of the UNDAFs under evaluation	11
Figure 11	→	UN-SWAP EPI Ranking System	14
Figure 12	→	UN-SWAP EPI ranking system per criterion	14
Figure 13	→	Guiding questions for each of the three UN SWAP Criterion	15
Figure 14	→	Breakdown of UN-SWAP scores for the 50 reports	15
Figure 15	→	Overview of the distribution of evaluations by region	16
Figure 16	→	Average score per region for the three criteria	16
Figure 17	→	Percentage of UNDAF Evaluations according to UN-SWAP EPI by region	17
Figure 18	→	Criterion One of the UN-SWAP EPI	17
Figure 19	→	Criterion One of the UN-SWAP EPI	19
Figure 20	→	No of males/females leads according to the UN-SWAP EPI	20
Figure 21	→	UN-SWAP Criterion One – Overall Scope and Evaluation Questions	21
Figure 22	→	Findings under Criterion One	21
Figure 23	→	Criterion Two – Guiding Questions	22
Figure 24	→	UN-SWAP Criterion Two	24
Figure 25	→	UN-SWAP Criterion Three – Findings Conclusions and Recommendations	25
Figure 26	→	Criterion Three – Findings	25
Figure 27	→	Criterion Three – context	26
Figure 28	→	<i>Criterion Three – Findings, conclusions and Recommendations</i>	28

Figure 29	→	GRES Scale	29
Figure 30	→	The percentage of gender references per strategic priority	29
Figure 31	→	Gender@Work Framework	30
Figure 32	→	Proportion of Overall Gender Changes Generated	31
Figure 33	→	Overview of Gender Changes generated and Gender Outcomes by Areas	32
Figure 34	→	Number of results and the changes according to the spheres of work under the Gender@Work Quadrant	33
Figure 35	→	Proportion of HR and GE outcomes and changes generated in the Area of Governance	34
Figure 36	→	Proportion of Gender Outcome and changes generated in the area of economic empowerment	35
Figure 37	→	Proportion of Gender Outcomes and Changes Generated in the Area of Environment and Climate change	37
Figure 38	→	Proportion of Gender Outcomes and Changes Generated in the Area of poverty	38
Figure 39	→	No of references to vulnerable groups in the evaluation reports	39
Figure 40	→	No of UNDAF Frameworks which reference indicators for vulnerable groups	40
Figure 41	→	Proportion of Outcomes Targeting vulnerable Group by Area	40
Figure 42	→	Overview of the type of GEWE and HR recommendations	42
Figure 43	→	Overall proportion of Gender and HR recommendations	43
Figure 44	→	UN SWAP 2.0 Indicator areas	43
Figure 45	→	Category of gender related recommendations	44
Figure 46	→	Overview of the type of gender related recommendations	45
Figure 47	→	No of references to disaggregated data in the evaluation reports	47

Introduction – Purpose and background

1.1 Background

The year 2015 marked the adoption of the 2030 Agenda for Sustainable Development with the Sustainable Development Goals (SDGs) being placed at the forefront of this commitment. The 2030 Agenda for Sustainable Development established at its core the concepts of "leaving no one behind" and "reaching the furthest behind first." This agenda addresses the structural causes of discrimination and inequality, including by focusing on those most at risk of being left behind. In addition to a standalone goal (SDG 5) on gender equality and women's empowerment (GEWE), gender dimensions are also embedded in specific targets in other goals covering a comprehensive set of issues, including the gender dimensions of poverty, hunger, health, education, water and sanitation, employment, cities and human settlements, climate change, access to justice and peaceful and inclusive societies.

Following the adoption of the SDGs, a new generation of United Nations Development Assistance Framework (UNDAFs) were developed to align with the 2030 Agenda. In 2019, the UNDAF was renamed the United Nations Sustainable Development Cooperation Framework (UNSDCF) ("Cooperation Framework") to more accurately reflect the contemporary relationship between governments and the United Nations Development System's (UNDS) collaboration to achieve the SDGs.²

Cooperation Frameworks will be perceived as the United Nations Office's principal mechanism for accountability to the Government. The Government's expectations of the United Nations Development System (UNDS) is therefore articulated through the new framework.

The framework is paramount in driving the United Nations Country Teams' (UNCTs) support to the country. "The country development programmes are derived from the UNSDCF, and not vice versa."³ The Cooperation Framework mandates UNCTs to comprehensively and adequately address, among others, human rights (HR) and gender equality (GE) concerns through clearly defined, gender-specific outcomes and outputs that contribute to relevant gender SDG indicators and targets, where appropriate.

To this end, evaluations of the UNDAF and future Cooperation Frameworks and the recommendations therefrom are paramount to ensuring that HR and GE are at the forefront of policy, and best practices and lessons learned are acknowledged and utilized for the future development of Cooperation Frameworks. HR and GE-responsive evaluations have gained prominence in recent years, and United Nations Evaluation Group (UNEG), among other partners, have developed a number of key guidance documents on how to ensure the mainstreaming of HR and GE into evaluation.

A number of key frameworks to assess and guide HR and GE mainstreaming have been developed in recent years. In 2012, in order to ensure greater UNDS accountability, a United Nations System-wide Action Plan (UN-SWAP) on GE was developed for tracking six main elements on gender mainstreaming: accountability, results-based management, oversight, human and financial resources, capacity and knowledge exchange and networking. A second generation of the UN-SWAP (2.0) was developed in 2018. The new framework extends the reach of UN-SWAP 1.0 by including new performance indicators on gender-related and SDG relevant results, as well as updating existing UN-SWAP 1.0 performance indicators.

² Given that this analysis was conducted on United Nations Development Assistance Framework evaluations published before the shift to Cooperation Frameworks, the term "UNDAF" is used throughout the report.

³ Adapted from the Draft of the *United Nations Sustainable Development Cooperation Framework (UNSDCF) Guidance, April 2019* (2019), p. 5.

Accordingly, the United Nations Country Team System-Wide Action Plan (UNCT-SWAP) Scorecard was also developed. The UNCT-SWAP Scorecard is an accountability framework at United Nations Country Team level which is a "standardized assessment of United Nations country-level gender mainstreaming practices and performance that is aimed at ensuring accountability of senior managers and improving UNCT performance." The UNCT-SWAP Scorecard was updated in 2018 in order to align with the SDGs and harmonized with the UN-SWAP to strengthen accountability for gender mainstreaming and development results at country level.⁴

A Scorecard for promoting and harmonizing reporting against the UN-SWAP evaluation performance indicator (EPI) that tracks performance on gender-responsive evaluation for all United Nations entities was also developed.

The EPI is "linked to meeting the gender-related UNEG Norms and Standards,⁵ which in 2016 recognized HR and GE as a standalone Norm."

Norm 8 – Human Rights and Gender Equality

The universally recognized values and principles of human rights and gender equality need to be integrated into all stages of an evaluation. It is the responsibility of 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".

In 2019, a record share of 70 percent of UNSCDFs featured GE among their primary objectives.⁶ GE is also the principal focus area of joint programs implemented by UNCTs (109 out of total 378). To date, however, there is little evidence on the extent to which the UNDAF evaluations that have been carried out are conducted in a gender-responsive manner and if and how they assess the progress of gender-specific results. To address this gap, this meta-synthesis of UNDAF evaluations with a GE lens was commissioned by the UNEG Working Group on Gender Equality and Human Rights.

1.1.1. Intended Use and Users of the Meta-Synthesis

The findings of this report are intended to serve as a reference to a number of potential users including the Resident Coordinators' Offices (RCOs), UNCTs, United Nations Gender Advisors and Focal Points, United Nations Regional Evaluation Groups, donor agencies, civil society organizations and national partners. Individual entities of the UNDS, UNEG and other inter-agency mechanisms are working on gender equality and the empowerment of women.

It is anticipated that every effort will be made to promote the findings of this report, particularly among the relevant policy and decision-making bodies, to potentially inform potential changes or amendments in relation to the subsequent development, implementation and evaluation of the future UNSCDFs. It is further envisioned that this report will also contribute to the conduct of a new generation of more gender-responsive UNSDCF evaluations and thus includes a specific set of recommendations to carry this forward.

Furthermore, the timing of the release of this report will be significant as 2020 approaches. The year 2020 will mark the twenty-fifth anniversary of the adoption of the Beijing Declaration and Platform for Action, as well as heralding a five-year milestone for achieving the SDGs of the 2030 Agenda for Sustainable Development.

⁴ Adapted from United Nations Sustainable Development Group, *UNCT-SWAP Gender Equality Scorecard – Performance Indicators for Gender Equality and the Empowerment of Women for United Nations Country Teams – 2018 Framework & Technical Guidance*.

⁵ United Nations Evaluation Group, *Norms and Standards for Evaluation*, June 2016 (2016), p. 12.

⁶ *Implementation of General Assembly Resolution 71/243 on the quadrennial comprehensive policy review of operational activities for development of the United Nations system (QCPR): Monitoring and reporting framework – 18 April 2019 (Add. 3)*.

1.1.2 Context

According to the 2019 UNDAF/UNSDCF guidelines, the conduct of an independent evaluation in the penultimate year of the UNDAF/UNSDCF period is a minimum requirement.⁷ UNDAF/UNSDCF evaluations assess whether the UNCT has contributed to transformative change that goes beyond the scope of programs and projects to help a country progress towards achieving the SDGs. Furthermore, the new guidelines require UNSDCF's evaluations to be conducted in an inclusive manner, meeting **the UN-SWAP EPI**, the evaluation indicator of the Inter-Agency Disability Inclusion Strategy, and to promote national ownership through **meaningful engagement of relevant national partners, and where possible, those representing the perspectives of marginalized groups**, throughout the evaluation process.

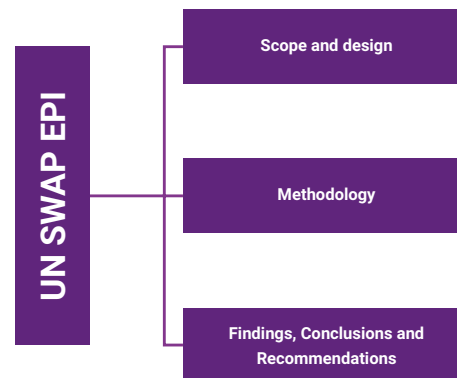
1.2 Purpose and Objectives of the Meta-Synthesis

The present meta-synthesis has two main objectives.

- 1. Assess the human rights and gender-responsiveness of a selection of UNDAF evaluations. This entailed, *inter alia*, an analysis of the extent to which the UNDAF evaluations were gender responsive, assessing the integration of human rights and the "leaving no one behind" principle using the UN-SWAP EPI criteria.**

Fifty UNDAF evaluation reports were sampled and assessed according to the three criteria under the UN-SWAP EPI. These examined the scope and design of the evaluation, the methodology used and the findings, conclusions and recommendations. The assessment examined the extent to which the evaluations have ensured that "human rights and gender equality values are respected, addressed and promoted, underpinning the commitment of the principle of 'no one left behind.'"⁸

Figure 1. UN-SWAP EPI criteria



- 2. Analyse the key findings and recommendations of UNDAF evaluations for gender equality results.**

The analysis concentrated both on HR and GE results reflected in the reports. The synthesis measured the extent to which these issues were incorporated and assessed the type of gendered change advanced. This was done by using the Gender Results Effectiveness Scale (GRES) and the Gender@Work Quadrant. (For more details, see section 2.1.3 and beyond.)

This report answers a number of questions and divides them into two components aligning with the two key objectives of the meta-synthesis: i) questions relating to the analysis of the evaluation reports against the UN-SWAP EPI and ii) questions relating to the analysis and synthesis of the individual reports and how they report on HR and GE-related results.

The questions are as follows:

First Component – Analysis of the evaluation reports against the UN-SWAP EPI

EQ1: To what extent did the UNCTs' commitment to HR and GE mainstreaming impact the quality of the evaluations and the extent to which HR and GE is mainstreamed in the evaluation reports?

7 United Nations Sustainable Development Group, *United Nations Sustainable Development Cooperation Framework – Guidance* (2018), p. 30.

8 United Nations Evaluation Group, *UN-SWAP Evaluation Performance Indicator Technical Note – Guidance Document – April 2018* (2018), p. 3.

EQ2: To what extent does a UNDAF which is rated as GE/HR-specific⁹ correlate with the extent to which HR and GE are integrated in the evaluations?

EQ3: What are the factors that can influence the extent to which HR and GE are integrated into the evaluations?

EQ4: To what extent were the evaluations under assessment found to be gender responsive?

Second Component – Analysis of the individual reports and how they report on HR and GE-related results

EQ5: What impact did gender equality, women’s empowerment and human rights interventions have?

EQ6: What have been the key drivers of change in gender equality and women’s empowerment?

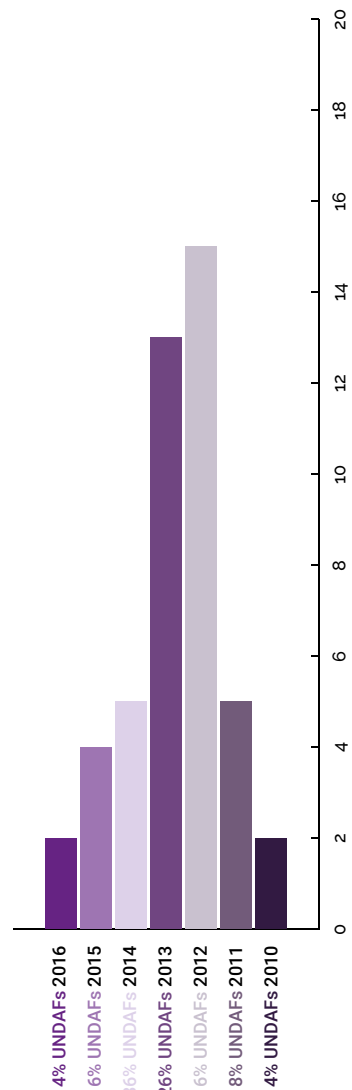
EQ7: To what extent have UNDAFs been able to address the leave no one behind principle?

EQ8: To what extent is inter-agency cooperation, including joint programmes and joint gender programmes, contributing to greater efficiency on HR and GE results, if at all?

1.2.1 Approach and Methodology

The meta-synthesis comprised an assessment of 50 selected UNDAF evaluations,¹⁰ which were conducted between 2015 and 2019.¹¹ The selected sample represented roll-outs from 2010 to 2016 (Figure 2). Accordingly, the design of the majority of the UNDAFs¹² under assessment were concluded prior to the introduction of the SDGs, and therefore the evaluations provide very little insight on the alignment to the SDGs and the leave no one behind principle.

Figure 2. Total number of UNDAF roll-outs included in the meta-synthesis



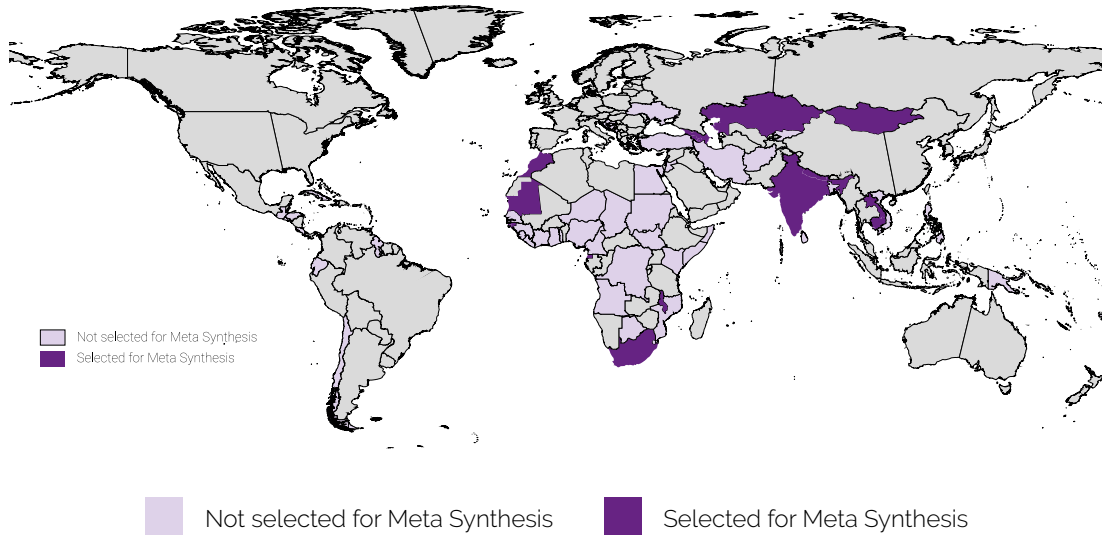
Nonetheless, the meta-synthesis has ensured that results deriving from the exercise reflect the extent to which these two principles are included and what will be needed in the future to enhance their inclusion.

9 The sample excluded mid-term reviews.

10 The original universe comprised 67 United Nations Development Assistance Framework (UNDAF) evaluations. Given the limited time and resources, fifty of these were selected for assessment. It should be noted that there is a possibility that more than 67 UNDAF evaluations exist, however the meta-synthesis was only able to obtain access to 67.

11 The sample includes one UNDAF from the 2016 roll-out and two from the 2015 roll-out.

Figure 3. Sample of UNDAF evaluations by geographic location



1.2.2 Sampling

Figure 4. Universal sample

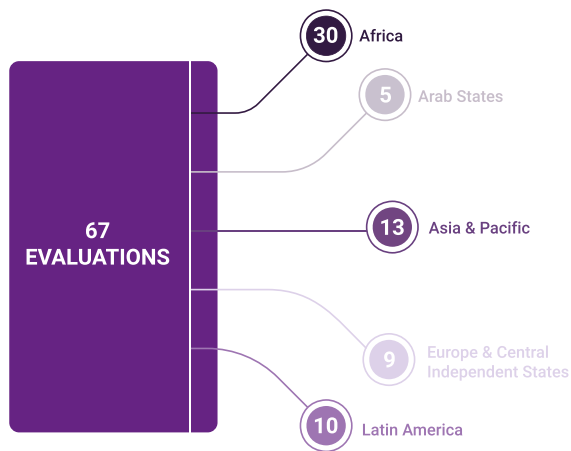
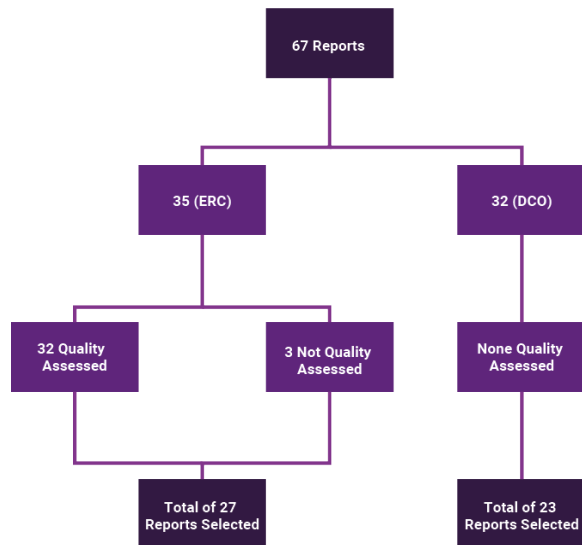


Figure 5. Overview of the reports in the universe



The meta-synthesis was able to access 67 reports in total. These were derived from two different sources. The preliminary list emanated from the United Nations Development Programme (UNDP) Evaluation Resource Center (ERC), which provided 35 reports that were in their repository. Thirty-two of these reports had been quality assessed.

The remaining 32 reports were sourced through the United Nations Development Coordination Office

(DCO) and had not undergone an external quality assessment, but were nonetheless included in the original universe. From the universe of 67 reports, a total of 50 reports were chosen through a lottery, irrespective of which list they had emanated from. Selected reports that had not been quality assessed subsequently underwent a “light” quality assessment using a standard evaluation quality assessment grid.¹³

13 The assessment was made based on the checklist of the Office of Internal Oversight Services (OIOS) for the assessment of UN Secretariat Evaluation Reports.

In total, 38 (76 percent) of the 50 sampled reports resulted in a quality assessment score of either highly satisfactory, satisfactory or moderately satisfactory. The remaining 12 evaluation reports (24 percent) were deemed lower in quality. In order to avoid a positive bias in the analysis, it was decided to maintain all the reports, irrespective of the score awarded to them during the quality assessments.

For the second component of the meta-synthesis it was envisaged that only higher quality evaluations would be included. However, it was concluded that the majority of reports did include references to gender results and consequently all 50 reports were included and assessed against each of the three UN-SWAP EPI, as well as assessments of the extent and quality of the reporting on gender and human rights results, according to the analytical frameworks described in the next section.

1.2.3 Approach to Meta-Synthesis

The principle method to carry out the meta-synthesis was a desk review of 50 randomly selected evaluation reports from a universe of 67. This was complemented by a vigorous desk review of relevant corpora relating to HR and GE-focused evaluations.¹⁴ The meta-synthesis also conducted interviews with six Resident Coordinators and two members of the DCO. Four interviews with evaluators were also held and a survey distributed to 20 persons with a response rate of 40 percent was also included in the analysis.

The meta-synthesis employed several variants of the UN-SWAP to analyse the evaluation reports. This included the UN-SWAP EPI which was used to determine the extent to which the evaluation reports

had integrated HR and GE in the scope and design of the evaluation, its methodology and the findings, conclusions and recommendations. A maximum score of nine could be allocated.

The second component of the meta-synthesis involved an in-depth analysis of the findings of each of the 50 reports. This included an examination of the recommendations. Due to time restraints, the analysis only examined the gender-related recommendations; this was carried out by utilizing the UN-SWAP II framework.

In order to determine the quality and content of the GE results included in each evaluation report, two analytical frameworks were used. This included the Gender Responsiveness Effectiveness Scale (GRES) which examined the effectiveness of each of the results and the Gender@Work Framework to determine how the UNDAF had contributed to the development results in the realm of HR and GE.

Given the limited time and resources, it was observed that the most common strategic priorities for the 50 reports under evaluation were the following focus areas: governance, economic empowerment, environment/climate change and poverty. Therefore, it was decided that the strategic priorities which covered these areas would be included in the meta-synthesis. In some cases, more than one of the focus areas was present in a single strategic priority, but each core focus area was recorded separately. The 50 evaluation reports had a total of 193 strategic priorities; 94 of these were examined. (See annex VIII for more details).

14

Please see annex for a full list of the documents consulted.

Figure 6. Methodological approach to meta-synthesis



1.3 Limitations

There are a number of limitations to this meta-synthesis. The universe of 67 reports was pulled from a number of sources; however, it is not known if there could be other UNDAF evaluations to which the meta-synthesis did not gain access. To this end, it is difficult to ascertain if the trends outlined in this report are fully conclusive or only represent part of the picture. In addition, as previously mentioned, all evaluations included in the final sample were subject to some form of quality assessment, either the UNDP version or a "light" assessment¹⁵ undertaken during the meta-synthesis itself. In total, nearly a quarter of the reports were deemed to be of poorer quality. Nevertheless, given the fact that all reports had integrated HR and GE to some degree and to ensure that there was no positive bias, it was decided to include all reports for the two exercises during the meta-synthesis.

Time and resource constraints prevented a more thorough analysis in some areas, including understanding the root causes of patterns that emerged during the meta-synthesis. Due to the fact that not all UNDAF evaluations appear to have been uploaded on the ERC website, it is impossible to know if each region under the meta-synthesis is adequately represented. To this end, it was decided to remove the key conclusions regarding regional trends as there was uncertainty regarding how representative this was.

It should also be noted that the meta-synthesis gained access to data of the budgets that were earmarked for 31 of the evaluations under the sample. However, the information was not used, as it could not be established whether this was the budget award or simply the provisional budget. A quick overview of the budget and possible trends did not bear fruit and it was therefore decided not to include the analysis on the budgets.

15

A light assessment was undertaken for all the reports selected, which had not previously undergone an official United Nations Development Programme quality assessment.

With regards to the analysis of the terms of reference (ToR), it should be noted that the meta-synthesis was able to gain access to 35 ToR out of the 50 evaluations. Therefore, some of the trends and commonalities encountered in the sample available were not comparable for the whole sample. Therefore, although conclusions were made regarding the significance of the ToR and the available ToR deemed to be representative of the overall sample, interpretations of results from this section should bear this in mind.

The meta-synthesis also envisaged the distribution of two surveys, one to the UNCTs and the other to the evaluators of the UNDAFs under assessment. The survey targeting the UNCTs was never distributed and the second survey of the evaluators was only successfully distributed to 20 evaluators as no official contact details were available and therefore it was difficult to find the correct email addresses. Given the timing of the distribution of the survey – which was sent out in the summer months – the response rate was quite low (40 percent) as only eight out of 20 evaluators (40 percent) responded to the survey.¹⁶ It was never foreseen in the original ToR to expand the scope beyond a desk review. To this end, the information emanating from the interviews and/or survey is anecdotal and by no means carries more weight than the findings of the desk review.

It is worth noting that the conclusions and recommendations contained in this report were derived from the information that was contained in the 50 evaluation reports. The information therefrom is by no means considered representative of the level and quality of results of the UNDAFs themselves.

1.4 Structure of the Report

The report has six main chapters. The first chapter is an introduction to the meta-synthesis, with an overview of its scope and methodology. The second

chapter addresses the first three questions outlined in the evaluation matrix. The first question examines the extent to which the UNCTs are an influencing factor in ensuring that evaluations are gender responsive.

The second chapter also looks at the impact a gender-responsive UNDAF can have on the integration of HR and GE and examines other possible factors that influence if and how HR and GE are integrated into the evaluations.

The third chapter then examines each of the 50 evaluations according to the UN-SWAP EPI. The use of this tool determined the extent to which each evaluation was deemed to be gender responsive. The chapter looks at the scope and design of the evaluations, the methodology undertaken and the findings, conclusions and recommendations of each evaluation report.

The fourth chapter then examines in greater depth the programmatic results recorded in the 50 evaluation reports in the sample. It assesses the extent to which the results integrate HR and GE perspectives using two key analytical frameworks, the GRES and the Gender@Work Framework. The analysis centers around four key focus areas; governance, economic empowerment, environment/climate change and poverty. This chapter also discusses the key challenges that were revealed by the desk review of the evaluation reports and consultations with key informant interviews (KII) on ensuring a HR and GE-responsive evaluation.

The fifth and final chapter makes a number of conclusions by summarizing the findings of the two components and puts forward recommendations for further action in order to promote the better integration of HR and GE in future Cooperation Framework evaluations.

16

There is no evidence that all twenty evaluators received the survey, as many of the contacts were reached via LinkedIn and Facebook and there was no guarantee that the evaluators accessed these accounts.

II. Key Findings - The Gender Responsiveness of UNDAF Evaluations

2.1 The Impact of Gender/Human Rights Specific¹⁷ UNDAFs on the Integration of Gender and Human Rights in the Evaluation Reports

In order to establish the correlation between gender-responsive evaluations and gender-responsive United Nations Development Assistance Frameworks (UNDAs), the meta-synthesis undertook a light assessment of the integration of human rights (HR) and gender equality (GE) in each of the UNDAFs being evaluated. The assessment adopted a similar methodology utilized in the desk review of 18 UNDAFs in Europe and Central Asia, making use of the same parameters (Figure 7).

Finding One:

A gender-specific/human rights UNDAF tended to ensure more responsive evaluations in terms of the integration of HR and GE perspectives.

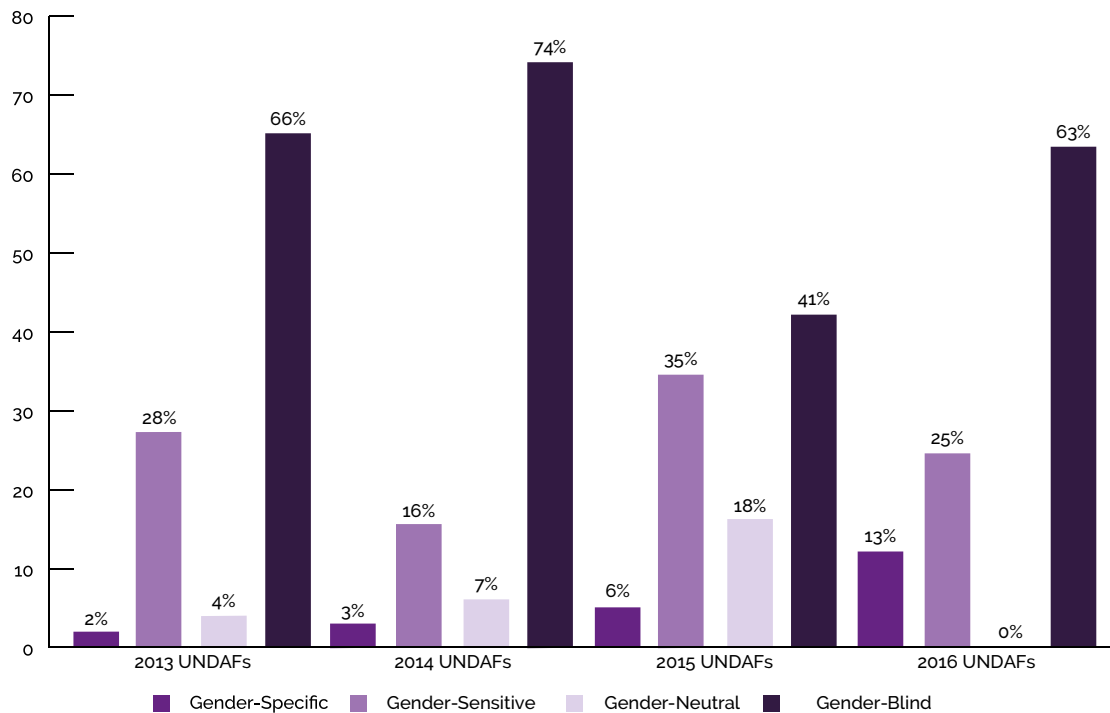
Over the last couple of years, the UNDAFs and/or Cooperation Frameworks are ensuring that the strategic priorities and/or outcomes are increasingly addressing inequalities and discrimination in order to ensure that no one is left behind. The number of gender-specific outcomes in the UNDAFs have increased every year and are currently at a rate of approximately 70 percent. Figure 8 substantiates that trend. The UNDAFs which were developed after 2015 showed a significant shift in the number of gender-specific outcomes in comparison to previous years.

Figure 7. Definition of outcome categories

Gender specific	Women, girls and/or related issues/needs are the focus of the outcome statement to promote gender equality and women's empowerment (gender responsive).
Gender sensitive	Women, girls and/or gender issues/needs are not the main focus of the outcome statement, but are mentioned in the outcome statement, footnotes or explanation in the narrative to contribute to gender equality and women's empowerment. An outcome with a focus on human rights can also be considered gender sensitive (gender responsive).
Gender neutral	An outcome statement that neutralizes any reference to gender, and thus cannot discriminate on the basis of gender.
Gender blind	An outcome statement that fails to recognize that gender is an essential determinant of social outcomes and does not consider women, girls and/or gender to be influential factors.

¹⁷ 'Gender specific' implies that women, girls and/or gender-related issues/needs are the focus of the outcome statement to promote gender equality and women's empowerment.

Figure 8. Percentage of UNDAFs deemed to be gender specific and gender sensitive



There are two types of interventions that are considered gender responsive: ones that are gender specific, i.e. target gender equality and women’s empowerment (GEWE) directly, and those that are gender sensitive, interventions which have the potential to contribute to GEWE.

Nonetheless, 29 percent are considered gender sensitive. To this end, out of a possible 194 strategic priorities for the 50 UNDAFs under evaluation, 65 were found to be either gender sensitive or gender specific.

When examining the number of specific priorities of all the UNDAFs under evaluation, on average, only 4 percent of the priorities are gender specific.

Figure 9. Percentage of gender-responsive priority areas of the UNDAFs under evaluation

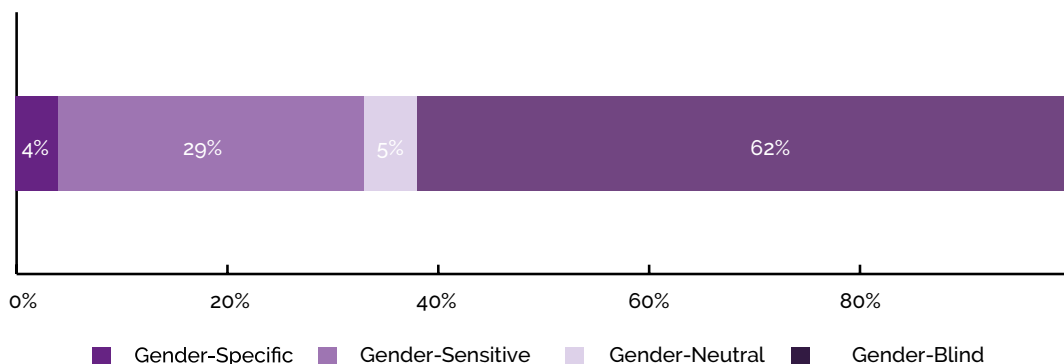
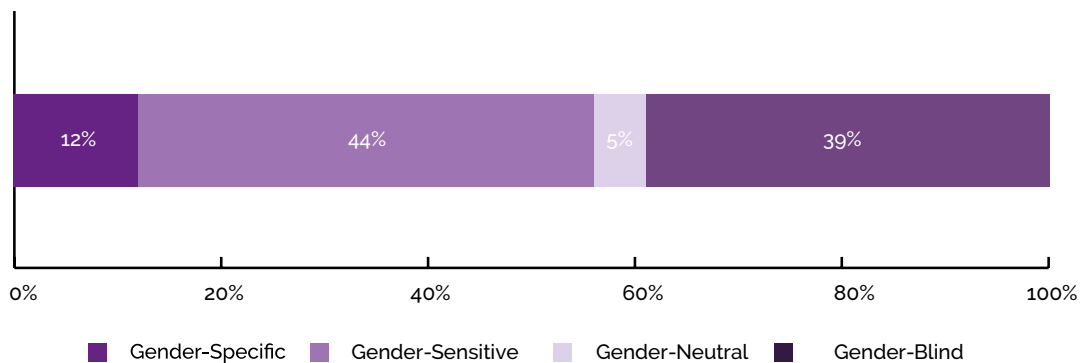


Figure 10. Percentage of gender-responsive outcomes of the UNDAFs under evaluation



The picture is far more positive at the level of outcomes, where in total, 66 percent of the outcomes are considered gender responsive (12 percent gender specific and 44 percent gender sensitive). This amounted to 295 of the 450 outcomes being gender responsive.

When looking at the individual scores, the higher the percentage of outcomes that were HR and GE-focused, the higher the likelihood of the evaluation approaching or meeting the requirements of the UN-SWAP EPI.¹⁸

As is illustrated in Figure 10, 81 percent of the UNDAFs with 60 percent or more of their outcomes being deemed gender responsive either approached or met requirements. Looking at the UNDAFs which missed requirements and had over 40 percent of their outcomes deemed gender responsive, 7 out of 11 of these were deemed to be poor-quality evaluations which might account for them missing requirements, despite being comprised of a significant number of gender-specific/sensitive outcomes.

The correlation to those that approached or met requirements and the number of gender-responsive

outcomes between 40 and 60 percent of outcomes was less pertinent, with only 30 percent approaching or meeting requirements. Nonetheless, this category had the highest percentage of met requirements. The two reports that met requirements, however, were also deemed to be of the highest quality and therefore the quality of the evaluators and the reports had an overwhelming impact on the integration of gender responsiveness in the reports in these two particular cases.

To this end, these findings substantiate, to a certain extent, that the degree to which an UNDAF is deemed gender responsive is one variable which could determine if an evaluation would be gender responsive or not. Nonetheless, as will be further discussed under chapter three, the quality of the evaluation is an important factor.

2.2 United Nations Country Team Commitments to Gender and Human Rights

The 2010 UNDAF guidelines¹⁹ state that the Country Teams are required to apply the five programming principles, two of which include a human rights-based approach and gender equality.

¹⁸ There were a total of 22 reports which had 0-30 percent of gender equality(GE)/human rights (HR) outcomes. Eleven of these missed, eight were approaching and three of them met requirements. There were 12 reports with between 31 and 40 percent GE and/or HR outcomes, with five *missing* requirements, six *approaching* and one *meeting* requirements. There were nine reports with between 40 and 60 percent GE and/or HR-focused outcomes. One *missed* and eight were *approaching* requirements. The remaining reports which had over 60 percent GE and/or HR-focused outcomes all *approached* requirements.

¹⁹ United Nations, *How to Prepare an UNDAF: Part (I) Guidelines for UN Country Teams, January 2010* (United Nations Development Group, 2010). See also: Part (II) Technical Guidance for UN Country Teams.

To this end, UNDAFs have become increasingly more HR and GE-specific and all United Nations agencies continue to promote both concepts in all their programmes.

Finding Two:

UNCTs are largely perceived as making concerted efforts towards integrating gender and human rights in many of the aspects of the UNDAF process, although some reports pointed to a number of weaknesses.²⁰

While the caliber of the evaluation team has an important bearing on the overall quality and the inclusion of HR and GE perspectives in an evaluation, there are a number of other factors which have the potential to influence the overall quality of an evaluation.

The UNEG guidance on integrating gender equality and human rights in evaluations²¹ outlines a framework which aims to promote accountability, learning and decision-making on HR and GE at all levels of an organization in relation to evaluations.

Nonetheless, while the framework is clear, the cost, time and human resources invested have the potential to impact the quality of the evaluation (including its GE and/or HR-responsiveness).

The cost of the evaluations is an important element. While the meta-synthesis could not draw any key conclusions from the budgets as only a third of the budgets were accessible, it is evident that not enough money is being set aside for the evaluations of UNDAF. Guidelines recommend that approximately \$100,000 should be accounted for, for an evaluation of this genre, but only four of the evaluations²² had over \$100,000 allocated to them. Furthermore, the average time allotted for conducting an evaluation tended to be around 25

days, leaving very little time for evaluators to be in the field in order to gather sufficient information based on evidence and draw judgments on the DAC criteria.

Perhaps the most significant weakness which inevitably impacts the quality of an evaluation is the effectiveness of the Monitoring and Evaluation Framework. Only two of the 50 reports referred to an effective monitoring and evaluation system.

All of the other 48 reports stated that the Monitoring and Evaluation Framework was either weak or needed strengthening. One of the key weaknesses was the quality of the indicators.

The quality of the design of the UNDAF, particularly of its indicators, undoubtedly had a negative effect on the quality of the evaluation.

Five out of the six Resident Coordinators' Offices (RCOs) consulted agreed that although HR and GE are indeed a priority for all United Nations offices, the practice sometimes fell short of the rhetoric. Reference was made to the fact that the RCOs do not always have the time or expertise to ensure gender and human rights mainstreaming across the whole of the UNDAF; these weaknesses were particularly pertinent when developing the results framework. The results framework and the annual progress reports, as well as Monitoring and Evaluation Framework are the key documents to which evaluators would then subsequently hold the Resident Coordinators to account.

According to three of the Resident Coordinators consulted,²³ and as cited in seven of the reports, the development of a results-oriented framework was often impeded by a lack of a theory of change and an overall vision as to what the UNDAF wanted to achieve. More often than not, although 20 out of the 50 evaluation reports stated that civil society was consulted in the design of the UNDAF, only three

20 The desk review revealed that United Nations Country Teams (UNCTs) generally demonstrated a commitment to gender equality and human rights. However, poorly designed frameworks, a scarce amount of gender and human rights-sensitive indicators, weak reporting and the failure to utilize specific working groups to their maximum sometimes weakened the position of the UNCTs. In some countries, this weakness was compounded by the commitments of external factors such as the government or the context of the country.

21 United Nations Evaluation Group, *Guidance Document – Integrating Human Rights and Gender Equality in Evaluations*, 2014.

22 It should be noted that the meta-synthesis was only able to obtain data on budgets for 31 out of 50 reports. Therefore, there may be other evaluations that had a budget higher than \$100,000.

23 Three of the resident coordinators spoken to mentioned this in the interview.

reports mentioned that civil society was involved in the design of the results matrix or the Monitoring and Evaluation Framework. It was alluded to by the Resident Coordinators consulted that the UNDAF tended to be designed at the last moment, and to this end RCOs were not always able to include HR and GE to the extent required/desired.

What is more, five out of the six RCOs consulted mentioned that the level of importance and commitment attached to implementation of the UNDAF (especially in small offices) was not always possible and sometimes the different United Nations agencies struggled to share the burden of implementation and monitoring equally. While the majority of the reports (48 out of 50) reported weaknesses in the monitoring and evaluation system, Delivering as One offices tended to cite that monitoring and evaluation was weak due to the United Nations agencies failing to implement a common strategy.

All the evaluators consulted referred to weak or the lack of consistent monitoring of results that further exacerbated the evaluators' inability to examine the gender and human rights aspects of an UNDAF.

The weaknesses in the original design of the UNDAF and sometimes the apparent lack of commitment to ensuring the integration of HR and GE in the implementation and the subsequent monitoring of the UNDAF also have the potential to hinder the overall quality of the evaluation. This is further undermined by the very small budgets available to carry out the evaluation. It is hoped that some of these bottlenecks will be addressed (and remedied) with the reinvigorated Resident Coordinator System and the roll-out of new UNSDCF guidance and accompanying compendium.

Finding Three:

The Strategic Priorities under the UNDAF which were deemed gender specific generally had a sufficient number of gender-sensitive indicators under the UNDAF against which progress could be measured. Nonetheless, Strategic Priorities which

were not specific to gender were harder to assess, and gender-sensitive indicators under the results matrix were generally not included and therefore it was far more difficult to evaluate gender-related results under other thematic results, such as climate change, economy and security.

The Secretary-General's report on mainstreaming a gender perspective into all policies and programmes in the United Nations system stated that "a total of 65 percent of UNCTs met or exceed requirements for gender mainstreaming in all UNDAF outcomes. However, only 44 percent translated commitments for gender quality into indicators to measure change."

In addition, where indicators did exist, as noted in the Moldova Report, "there [is] a tendency to consider gender-sensitive programming from the perspectives of women only, addressing less the areas where males are disadvantaged (e.g. dropping out in secondary schools, men's health issues, etc.) [...] left aside or only marginally addressed the needs of several groups whose human rights are repeatedly violated (notably the Lesbian, Gay, Bisexual, Transgender, Queer and Questioning – LGBTQ+ – community, religious minorities, illegal migrants), although they have been figuring in the definition of vulnerable groups in the UNPF."²⁴

However, one evaluator did state that while the United Nations agencies are committed to examining HR and GE results, some of the key stakeholders – such as the Government, who is also one of the key owners of the evaluation reports – are more sensitive to such topics and therefore the evaluator faced the dilemma as to whether, and to what extent, to include an analysis to appease the evaluation's target audience. This is particularly salient with regards to both the inclusion of disaggregated data and results relating to other groups such as the Lesbian, Gay, Bisexual, Transgender, Queer and Questioning (LGBTQ+) community, as well as a number of ethnic minorities.

To this end, UNCTs face numerous challenges to ensure HR and GE-responsive evaluations.

24

United Nations Moldova, *Final Evaluation of UN – Republic of Moldova Partnership Framework 2013–2017 – August 2016* (2016), p. 45–46.

III. Key Findings - UN SWAP EPI

3.1. Assessment of the Gender Responsiveness of the Fifty UNDAF Evaluations

3.2. Methodology

Figure 11. UN-SWAP EPI ranking system



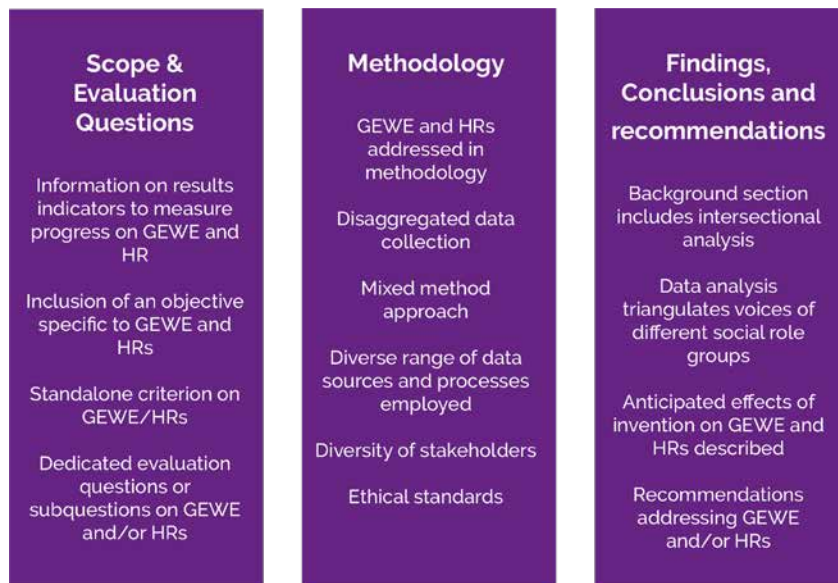
In order to determine the gender equality (GE) and human rights (HR)-responsiveness of each of the 50 selected evaluation reports, a desk review was conducted assessing the integration of HR and GE against the UN-SWAP EPI criteria. Each report was assessed as either missing, approaching or meeting requirements. A four-point scale (0-3) rating system was utilized to measure the extent that HR and GE were integrated under each criterion. A maximum of nine points per report could be assigned.

Figure 12. UN-SWAP EPI ranking system per criterion



The assessment examined (i) the scope and design of the evaluation; (ii) the methodology and the (iii) findings, conclusions and recommendations. A short summary of the findings was provided for each report, outlining which elements were included in the reports in accordance with the individual criterion.

Figure 13. Guiding questions for each of the three UN-SWAP criteria



3.2.1. Integration of Human Rights and Gender Equality Perspective

Finding Four:

Only four of the reports (8 percent) were considered to be gender responsive (met requirements) in terms of their integration of HR and GE, although another 27 reports (54 percent) were deemed to be partially gender responsive (approaching requirements on the UN-SWAP EPI Scorecard).

Figure 14. Breakdown of UN-SWAP scores for the 50 reports

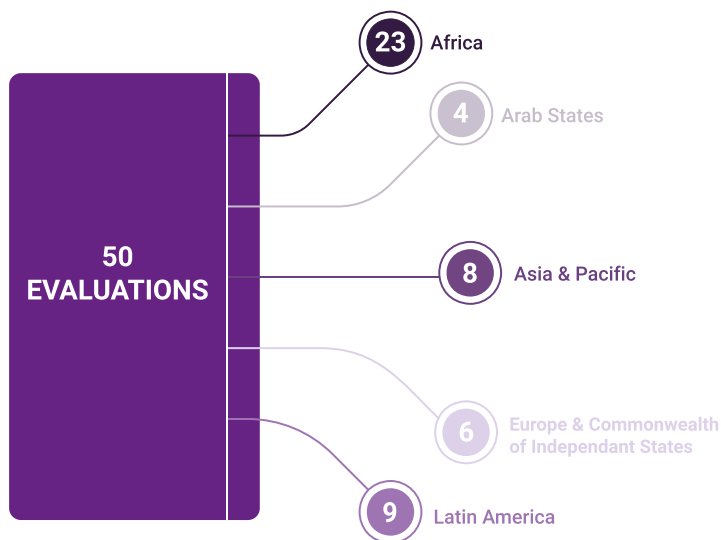


On average, the third criterion, which determined the extent to which the **findings, conclusions and recommendations** of the reports were found to be HR and GE-responsive, **scored considerably higher than the first two criterion.**²⁵

This is to say that the evaluation reports integrated HR and GE into the findings, conclusions and recommendations to a greater degree and with greater depth than was noted in the scope and/or methodology.

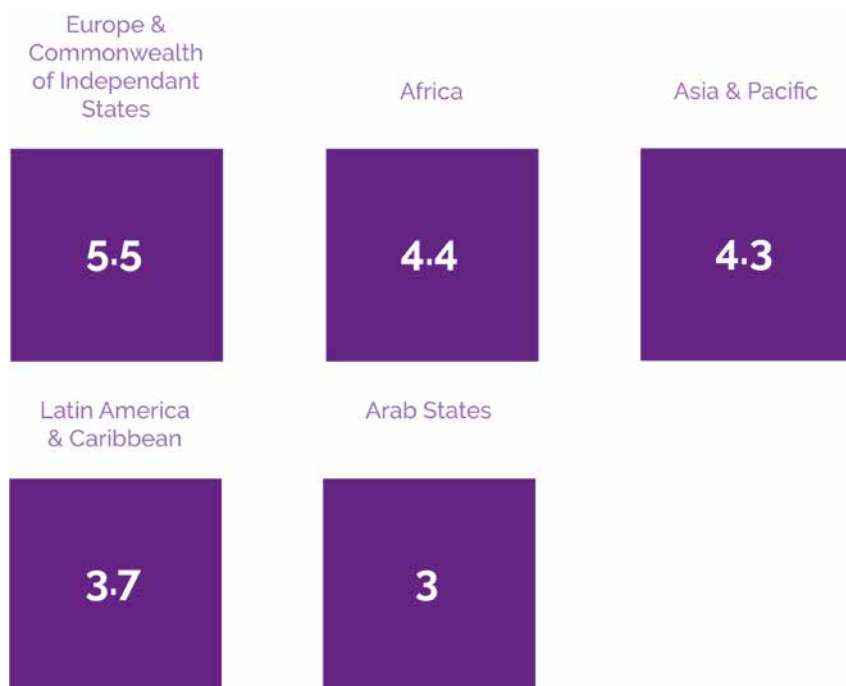
²⁵ On average, the first criterion scored 1.34, the second criterion 1.36 and the third criterion 1.56.

Figure 15. Overview of the distribution of evaluations by region



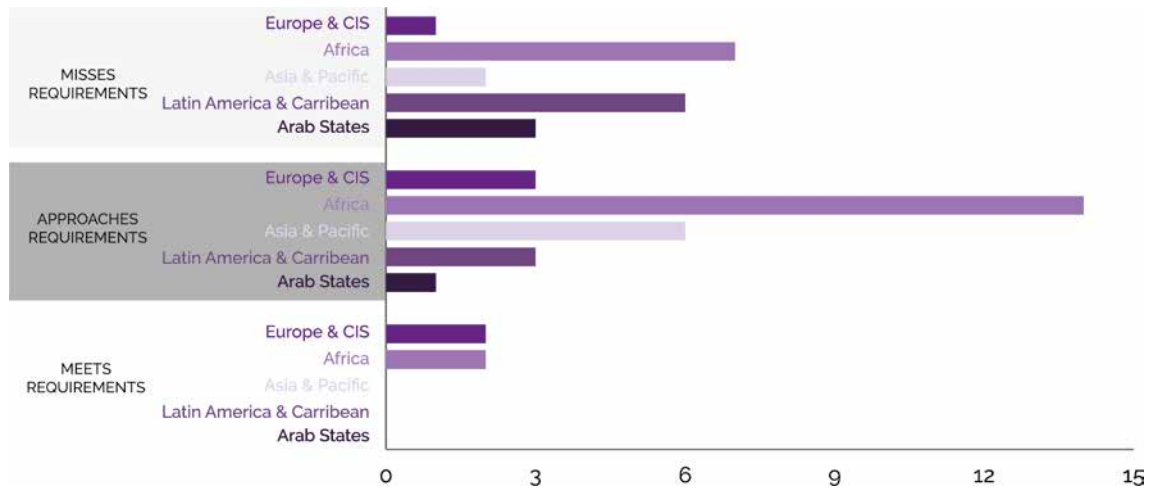
The assessment comprised of 50 evaluations representing evaluations from across five different regions. The distribution is representative of the original sample of reports provided for the meta-synthesis.

Figure 16. Average score per region for the three criteria



The variations of the results of the UN-SWAP Criteria EPI according to each region is quite pertinent in the sample. **The European and Commonwealth of Independent States (CIS) region had the highest average cumulative score (5.5), followed by Africa (4.4) and then the Asia and Pacific region (4.3).** All three regions had an average score which **approached requirements**. We can see from Figure 17 that in the Asia and Pacific region, 6 (71 percent of the total reports in the region) of the evaluations “*approached requirements*,” and in Africa, 16 (62 percent of the total reports in the region) either “*approached*” or “*met*” requirements. In the Europe and CIS region, which had the highest number of “*met requirements*” (in terms of percentages), 5 (83 percent of the total reports in the region) of the evaluations either “*approached*” or “*met requirements*.”

Figure 17. Percentage of UNDAF evaluations according to UN-SWAP EPIs by region

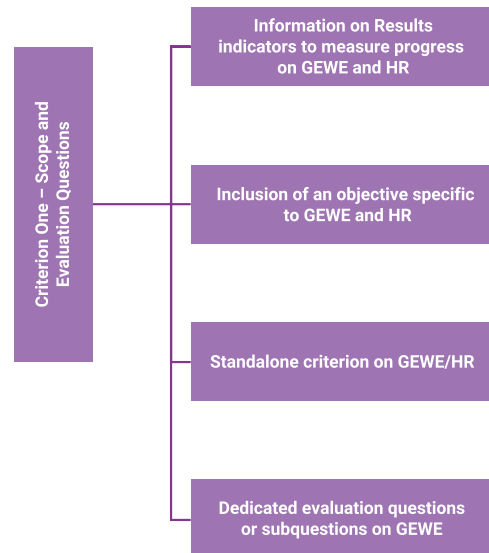


On the contrary, the results of the Arab region and Latin America and the Caribbean are particularly low, with both regions scoring an average of 3.7 and 3.0 points respectively, resulting in both regions **missing requirements**. Both regions fail to deliver any evaluation which met requirements, and in percentage terms, both regions have more missed requirements than approached requirements.

The reason for the weaknesses in these two regions is unclear and will inevitably require further analysis. As can be seen in the next section, the overall quality of the reports did correlate with the extent to which the reports incorporated HR and GE perspectives. As is illustrated in Figure 17, both the Arab States and Latin America and Caribbean had more reports which were deemed lower quality than reports from the other three regions.

Criterion One – Scope and design

Figure 18. Criterion One of the UN-SWAP EPI



The level of integration of HR and GE into the scope of analysis and evaluation criteria differed enormously between the reports. In some cases, the description of the scope mirrored the ToR, and therefore where a GE or HR scope was not included in the original ToR, evaluators tended to not also include a GE and/or HR approach.

Nonetheless, there were a few examples of reports which actually added such a component despite the absence of one in the ToR.

3.3 The Impact of the Terms of Reference on UNDAF Evaluations

Finding Seven:

HR and GE-focused ToR are important to promote the incorporation of a HR and GE perspective in the final evaluation.

The terms of reference (ToR) are key in ensuring that GE and/or HR perspectives are explicit and thus has the potential to bring these issues to the “front and center of the entire evaluation process.”²⁶

In total, 35 UNDAF evaluation ToR²⁷ were assessed against the criteria established as good practice in the United Nations Evaluation Group (UNEG) guidance document on preparing the evaluation terms of reference when integrating human rights and gender equality in evaluations.²⁸

The analysis looked at the extent to which the scope of the evaluation was sensitive to HR and GE, whether evaluation questions on HR and GE were included and whether evaluators with a background and/or expertise in HR and GE was requested.

Taking the sample of the 35 ToR made available to the meta-synthesis, the results demonstrate that there is a correlation between the inclusion of some of these good practices and the overall score on the UN-SWAP EPI. This was particularly pertinent with regards to the inclusion of HR and GE within the overall scope of the evaluation, as well as the methodology.

Twenty-one out of the 35 (60 percent) ToR in the sample failed to include either a GE and/or HR-responsive scope. Of the 14 (40 percent) ToR which included a GE and/or HR-responsive scope, two of them missed requirements.

This included Central African Republic and Costa Rica, both of which were quality assessed as poor.

Nineteen ToR (54 percent) included at least one question on GE and/or HR, five of these reports missed requirements.²⁹ Three out of the five were poorly assessed for quality.

A total of 20 (57 percent) ToR requested some type of experience or expertise in GE and/or HR.³⁰ Two of the evaluation reports missed requirements. Both reports were deemed unsatisfactory in the quality assessment.

In theory, ToR have the potential to guide the evaluation. However, some of the evaluators consulted highlighted weaknesses in the ToR they had received. Each evaluator/evaluation team is asked at the beginning of their assignment to outline the methodology and their evaluation questions in an inception report. These normally fall in line with the scope and methodology outlined in the ToR, albeit that the evaluator(s) in essence are at liberty to expand upon or otherwise change or tweak the evaluation questions and the approach they use to carry out the evaluation. Should these changes envisage a weaker HR and GE-responsive evaluation than planned, the UNCT has the opportunity to intervene and question the choices of the evaluator(s) if, for example, the evaluator(s) have failed to include the same range of scope as envisaged in the original ToR and thus fails to consider a GE and/or HR perspective when these were in fact highlighted in the ToR.

26 United Nations Evaluation Group, *Guidance Document – Integrating Human Rights and Gender Equality in Evaluations* (2014).

27 Seventeen in English, and six each in French and Spanish. Fifteen of the reports came from Africa, two from Asia and Pacific, five from Europe and the Commonwealth of Independent States and seven from Latin America and the Caribbean.

28 United Nations Evaluation Group, *Guidance Document – Integrating Human Rights and Gender Equality in Evaluations* (2014), p. 60.

29 These included the Central African Republic, Chile, El Salvador, Guyana and Turkey.

30 It should be noted, however, that it was not possible to confirm who and what type of expertise each evaluator had. Therefore, there is currently no information on the extent of the terms of reference or which requested persons with expertise on gender equality and human rights were recruited. Furthermore, among the terms of reference which did not explicitly cite the need for expertise on gender equality and/or human rights, it is unknown if the actual evaluators selected had this expertise or not.

The evaluator(s) also have the freedom to include a GE and/or HR approach, even if it is not explicitly specified in the original ToR. In two of the four reports that “met requirements,” the original ToR did not include any GE and/or HR-sensitive guiding questions at all.

Nonetheless, the evaluators/evaluation teams designed their own questions, ensuring the inclusion of a wide range of GE and/or HR-related questions in the evaluation reports, despite there being no reference thereto in the ToR.

Given the tight timeframes and limited resources allocated to the conducting of an UNDAF/Cooperation Framework evaluation, it is essential that the exclusion of a gender-sensitive approach to evaluation be identified prior to data-collection and synthesis, as there is normally very little room for manoeuvre and change the direction of the approach or methodology once the evaluation has begun.

3.5 The Importance of Ensuring Evaluators with Human Rights and Gender Equality Experience

Finding Eight:

Mixed teams of international and national experts tended to produce better results, although having a gender-balanced team did not necessarily guarantee a gender-responsive evaluation.

When undertaking a HR and GE-responsive evaluation, it is best practice to have a gender-balanced team. However, as the results will illustrate, **having a gender-balanced team did not always guarantee a gender-responsive evaluation.**

Nonetheless the **analysis concluded that a mixed team of international and national expertise tended to lead to evaluations which were deemed to be more gender responsive. Equally, the analysis revealed the importance of possessing intersectoral diversity (i.e. on climate change and governance, among other sectors) and someone who is versed in HR and GE in the United Nations context as well as the national context.**

Four of the RCOs which were consulted, however, highlighted that it was often very difficult to get an experienced evaluator with HR and GE expertise. Only 20 out of 35 (57 percent) ToR highlighted the need for expertise in these areas. Furthermore, while a possible candidate may demonstrate HR and GE experience in their CV, the quality of the evaluators is not always guaranteed. It is therefore important to ensure that all evaluators and their work are quality assessed. It is vitally important to share the overall assessment of the evaluators on a combined roster or a specific portal in order to ensure that country offices are able to promote the hiring of quality evaluators and slowly weed out those evaluators who have demonstrated less competence in the area of evaluation and/or HR and GE.

Figure 19. Number of male/female leads, according to the UN-SWAP EPI



(Currently, there is no official oversight to ensure the blacklisting of incompetent evaluators or provide overall feedback to the evaluators on the quality of their work.) To this end, offices often find themselves selecting poor-quality applicants and can do little to really improve the quality of the evaluation thereafter.

A total of four evaluations were led by an external agency which contracted individuals to carry out the evaluations. In most cases, the names and/or the gender of the members of the team were not included, making it quite difficult to gauge the gender of each of the selected evaluators. The use of an external agency was not a guarantee to ensure a gender-responsive evaluation (50 percent of evaluations overseen by agencies missed requirements).

With regard to the four evaluations which met requirements, the evaluations were carried out by an equal number of male and female lead evaluators. It should also be noted that all four evaluations were conducted by teams of at least two people, (one international and one national) and all teams had both female and male evaluators.

Finding Nine:

The meta-synthesis found that the integration of HR and GE into the scope and overall design of the questions in the evaluation reports was relatively weak. Nevertheless, the lack of reference to a GE or HR scope did not necessarily mean that the analysis in the findings did not integrate a HR and GE approach.

The first criterion of the UN-SWAP EPI – focused on the scope of the evaluation and evaluation questions – was found to be the weakest of the three criteria – was found to be the weakest of the three criteria with 30 (60 percent) of the reports either scoring one (partially integrated)³¹ or 0 (not at all integrated)³² points.

Scope and evaluation questions integrate human rights and gender equality

Figure 20. UN-SWAP Criterion One – Overall scope and evaluation questions



³¹ Applies when some minimal elements are met but further progress is needed and remedial action to meet the standard is required.

³² Applies when none of the elements under a criterion are met.

Dissecting each guiding question under Criterion One also revealed some significant results. **Only 19 reports (38 percent) integrated HR and GE in the evaluation scope.** Nonetheless, 34 reports (68 percent) included at least one question related to HR and GE under one or more of the DAC evaluation criteria. All 19 reports that had integrated HR and GE in the evaluation scope introduced at least one GE and/or HR-related question in the evaluation matrix, thus suggesting a positive correlation.

Furthermore, and more importantly, **all the evaluations which either satisfactorily or completely integrated the elements under this criterion approached or met requirements.**

Scope and evaluation questions integrate human rights and gender equality

Figure 21. Findings under Criterion One



Finding Ten:

Inclusion of HR and GE questions in the design of the evaluation tended to lead to a more gender-responsive evaluation. Questions were usually included under the criterion of relevance and/or effectiveness. A handful of evaluations had mainstreamed GE and/or HR across all criteria upon their own initiative.

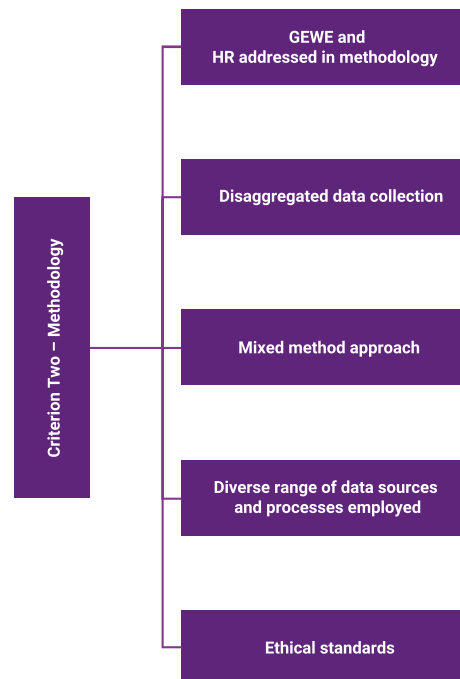
The reports which included HR and GE-responsive questions normally scored quite high on the UN-SWAP EPI overall, and either approached requirements or in the case of four reports, met requirements. Twenty-two out of the 50 reports included a subsection/section on GE and/or HR, but many of these sections generally focused on the status of women and did not include information on

the results in relation to GE and/or HR.

In total, only seven of the reports included a section on GE and/or HR within the key results under the relevance and/or the effectiveness section. Only four of the reports³³ included an independent section dedicated to either GE and/or HR outside of the four or five DAC criteria and/or the five programme principles.

Criterion Two – Methodology

Figure 22. Criterion Two – Guiding questions



The second criterion, which examines the extent to which a gender-responsive methodology, methods, tools and data analysis techniques are selected was also found to be weak, with 30 reports (60 percent) either not integrating any aspect of GE and/or HR at all or only in part.

This particular criterion was quite difficult to assess, as it was – at times – challenging to ascertain whether certain aspects of the cited methodology had actually been implemented. Furthermore, it was assumed that the evaluators had not always

33 Cameroon, Guinea, Niger and Senegal.

provided all the details of their methodology. Therefore, a GE and/or HR-sensitive methodology may have been used but this was not explicitly detailed in the approach cited.³⁴ Three of the reports annexed the methodology in order to outline in greater detail the approach and tools they had used.

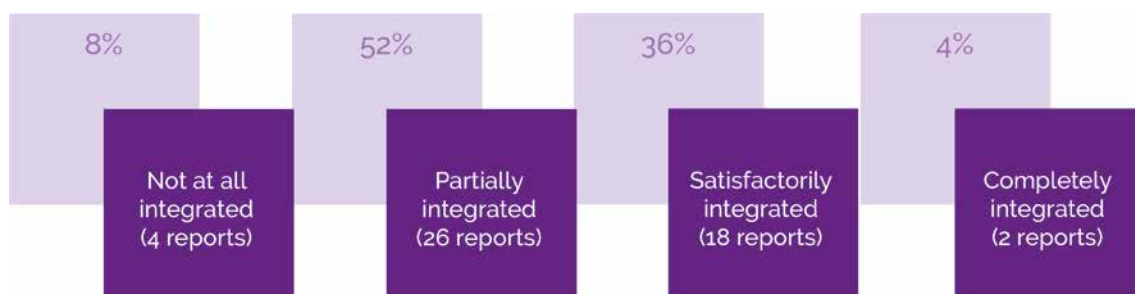
Approximately a fifth of the reports (20 percent) cited that the methodology was in line with UNEG guidelines in general, but none referred to the UNEG guidance on the integration of HR and GE.

In general, a mixed-method approach was cited as being utilized (three reports did not cite a methodology at all or used a different

approach). Furthermore, the majority of the reports demonstrated that a diverse range of data sources and processes are being employed. Nonetheless, many of the evaluations did not refer to a wide range of stakeholders being consulted, with the majority not interviewing any vulnerable groups or women-only groups (see finding twelve and beyond). Furthermore, although about a fifth of evaluators implied the use of ethical standards by citing compliance of the evaluation reports with the UNEG guidelines, **ethical standards were only explicitly mentioned in 11 (22 percent) of the reports.**

Methodology and tools integrate human rights and gender equality

Figure 23. UN-SWAP Criterion Two



Finding Eleven:

Only 11 (22 percent) reports included disaggregated data, with much of this data being cited in the background section of the report. Nonetheless, it was stated in approximately 75 percent of the reports that evaluators had encountered difficulties in collecting data and finding accurate information on the achievement of the targets outlined in the indicators. These difficulties were not just restricted to HR and GE results, but was a problem overall.

The evaluators consulted (both in the surveys and in the interviews) stated that there was a lack of HR and GE-sensitive indicators against which data could be collected and progress monitored, and when such indicators did exist, they tended to be gender blind or at best gender targeted.

Nonetheless, only three evaluation reports³⁵ explicitly mentioned the lack of gender-sensitive indicators.

³⁴ It is not all together clear whether the failure to cite a gender equality or human rights methodology was due to it not being in the terms of reference (the meta-synthesis was only able to gain access to 35 of the terms of reference).

³⁵ Cameroon, Guatemala and Moldova.

The evaluation report from Cameroon analysed all the indicators under the UNDAF and concluded that only 23 percent of the indicators were gender sensitive. The lack of HR and GE-specific indicators, although not cited as a problem in the majority of the reports, probably presents one of the key challenges to integrating a HR and GE approach to the evaluation. While the meta-synthesis was unable to conduct an in-depth analysis of the availability of HR and GE-sensitive indicators in the UNDAF results matrix, a quick scan of the UNDAFs tended to indicate that GE-sensitive indicators were only present under outcomes which specifically targeted women and/or girls. Other types of indicators to measure HR were rarely used.

Finding Twelve:

Only three reports mentioned consulting individuals outside of the "key stakeholders" in the methodology.³⁶ When looking at the list of the persons consulted in each evaluation report, five reports out of 50 mentioned meetings with other groups, such as women's groups, different representatives from a number of institutions, a refugee group and a civil society organization representing persons with disabilities.

The evaluators who were consulted and/or responded to the surveys stated that it was not always possible to consult with women's groups, communities and other vulnerable groups. It was unclear if the request to meet with these groups was highlighted in all the corresponding ToR or not.³⁷

The reasons for not interviewing other groups varied and included a lack of access to these communities, a lack of resources and a lack of time to contact these groups.

There was also concern as to how representative a small group would be with regards to the UNDAF and if indeed the group in question would understand what the UNDAF is and what impact it would have on their programme.

This led to evaluators having to prioritize interviews with the key stakeholders such as the Government and the ministries, as well as project staff and United Nations agencies. The survey revealed that only two out of eight of the evaluators felt that they had been able to adequately consult with rights holders during the evaluation, while six respondents either considered that the consultation was "somewhat adequate" or "inadequate."

Traditionally, in order to ensure a gender-responsive evaluation, the central focus of the evaluation should look upon the gender inequities that lead to social injustice. It is also important to assess whether discrimination or inequality is systemic or structural. There are a number of tools which are traditionally used for gender-responsive evaluations. These include outcome harvesting, which looks to "identify, describe, verify and analyse outcomes."³⁸ Other tools include the most significant change, which "involves generating and analysing personal accounts of change and deciding which of these accounts is the most significant – and why."³⁹

" Such tools were rarely used and a feminist approach⁴⁰ to evaluation was rarely evident. The majority of the evaluations used data-collection tools such as key informant interviews (KII), surveys and desk reviews. Feedback from surveys was not always sufficient and on a few occasions responses from the surveys were attributed more weight in the analysis, as they could not obtain the required data through the documentation (i.e. indicators) due to their unavailability or the collection not being done in the first place.

³⁶ Although when consulting the list of the persons consulted, reference to these groups was not found.

³⁷ The meta-synthesis was only able to obtain copies of the terms of reference for 35 of the 50 evaluations. About one third of the terms of reference referred to meeting other groups such as women's groups or civil society organizations.

³⁸ INTRAC, "Outcome harvesting", 2017. Available at <https://www.intrac.org/wp-content/uploads/2017/01/Outcome-harvesting.pdf>

³⁹ BetterEvaluation, "Most Significant Change", n.d. Available at https://www.betterevaluation.org/en/plan/approach/most_significant_change.

⁴⁰ Addresses the gender inequalities that lead to social injustice and examines opportunities for reversing gender inequities – see *United Nations Evaluation Group, Guidance Document – Integrating Human Rights and Gender Equality in Evaluations*, 2014.

It was often the case that either the UNCT and/or the Government **did not collect or was not able to collate all the information on the indicators** and this sometimes led to the evaluators having to look to other resources to draw their conclusions.

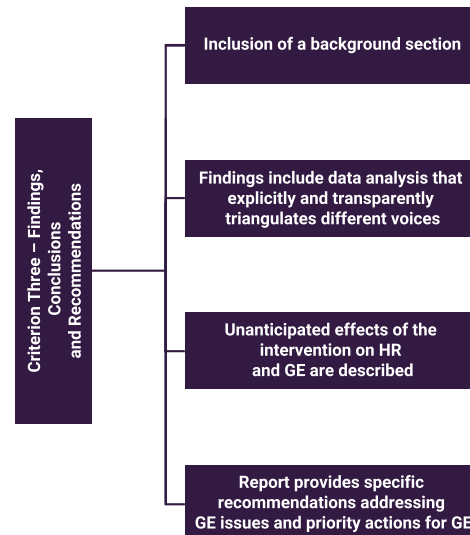
This difficulty was exacerbated by the high number of indicators included in many of the results matrices. In addition, on many occasions, the indicators were not considered appropriate and/or SMART.⁴¹ Annex VII illustrates the range of the number indicators for the 50 UNDAFs under evaluation.

These ranged from 12 to 348 indicators. To this end, many of the reports struggled to demonstrate results at an outcome level. The lack of information resulted in some of the evaluations reporting on the number of women targeted under the interventions rather than how they had benefited from the intervention. The lack of information was highlighted by the Papua New Guinea report,⁴² which stated:

“ The issue is not that there is no impact, the issue is that we are unable to demonstrate impact.” The ability to measure the impact of the UNDAF at both the outcome and output level is weakened due to the poor quality of the UNDAF Results Framework and a limited evidence base from which to draw data so as to measure and attribute results.”

Criterion three – Findings, conclusions and recommendations

Figure 24. UN-SWAP Criterion Three – Findings, conclusions and recommendations



Finding Thirteen:

Evaluation reports generally scored higher under Criterion Three, which examines the extent to which they integrate HR and GE in the findings, conclusions and recommendations. This finding correlates with the fact that 66 percent of the UNDAF outcomes in the results matrix were deemed to be gender responsive and therefore information on gender results would theoretically have been available. Nonetheless, references to HR and GE-responsive results were scattered across the evaluation reports and did not entirely correlate with the high number of gender-responsive outcomes from the UNDAF Results Framework.

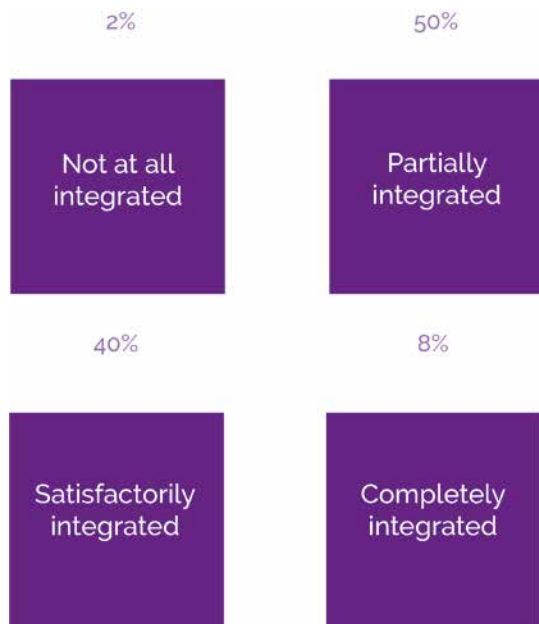
⁴¹ Specific, Measurable, Attainable, Relevant and Timebound.

⁴² United Nations in Papua New Guinea, *Independent Evaluation – UN Development Assistance Framework (2012-2017) UN Country Team in Papua New Guinea – June 2016* (2016), p. 24.

In total, 66 percent of the UNDAF outcomes were classified as gender responsive. While the degree to which an individual UNDAF was gender responsive did affect the degree to which the evaluation had integrated HR and GE, the evaluations still fell short of ensuring that the entire evaluation was gender responsive. Most of the evaluators limited themselves to reporting on a few results in a gender-responsive way and did not see the entire evaluation and its results through a gender and/or human rights lens. To this end, the references to HR and GE results reported and highlighted in the evaluation reports represented a very small proportion of the entire report.

Findings, conclusions and recommendations integrate human rights and gender equality

Figure 25. Criterion Three – Findings



Finding Fourteen:

A total of 31 reports (62 percent) contained some reference to GE and/or HR in the background section. A comprehensive background section including references to the HR situation, reference to compliance of international GE or HR instruments and the mentioning of how the intervention impacts on both genders was found to be included in 12 of the 31 reports (38 percent) which had cited GE and/or HR in the background section.

Inclusion of human rights and/or gender equality perspective in the background section of the report

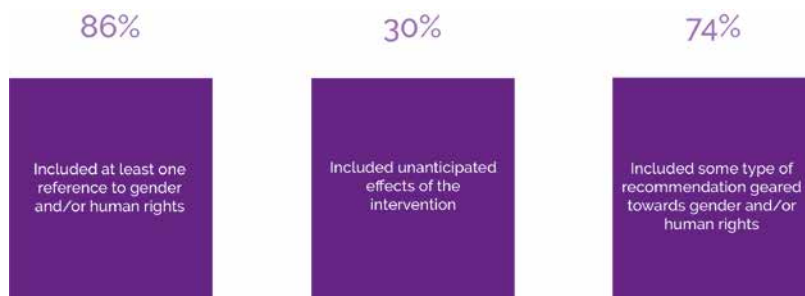
Figure 26. Criterion Three – Context



The evaluator’s survey confirmed this finding, whereby only one respondent stated that he/she had adequately included a gender analysis in the context section. All the others stated that they had either included it somewhat “adequately” or “inadequately.” All eight respondents to the survey concluded that they had not adequately included an analysis of the extent to which internationally and nationally agreed norms on HR and GE are met by support under the UNDAF.

Criterion Three – Findings, conclusions and recommendations

Figure 27. Criterion Three – Findings, conclusions and recommendations



While the analysis of how HR and GE results were captured in the evaluation reports differed from one report to another, 43 of the 50 reports (86 percent) included some reference to HR and GE. However, the quality and extent to which the findings were HR and GE-responsive varied from one report to the next and was dependent on the area that was being discussed (i.e. governance, environment/climate change, economic empowerment, poverty.)

A total of 15 (30 percent) noted unanticipated effect(s) of the intervention, and 34 (74 percent) reports included some type of recommendation(s) geared towards GE and/or HR. The types of recommendations differed; where some evaluators just stated that the importance of HR and GE should be increased in the upcoming UNDAF, others dissected the particular results, providing information on how they could have a more gender-responsive impact in future.⁴³

In the survey completed by the evaluators, only one respondent stated that they had “adequately” included HR and GE in their conclusions, and two stated that they had either “adequately” or “somewhat adequately” included gender in the recommendations. All the other respondents stated that they had “inadequately” provided recommendations on gender.

⁴³ For more information on the recommendations, see Chapter 4.

IV. Key Findings - Meta-Synthesis of Gender Results

4.1 Introduction

Under the second component, the meta-synthesis focuses on the extent to which human rights (HR) and gender equality (GE) are integrated in programmatic results and recommendations captured in the findings of the United Nations Development Assistance Framework (UNDAF) evaluations. The analysis will examine the extent to which these references have contributed to HR and GE results and whether they represent a transformative change. In order to complement the findings from the 50 evaluation reports in the sample, the analysis also utilizes anecdotal evidence from the consultations with the evaluators and key informant interviews (KII) with United Nations Country Teams (UNCTs), and the Development Coordination Office (DCO). However, the findings therefrom do not have the same weight as the desk review.

The analysis under this section will also examine the extent to which inter-agency cooperation contributed to greater efficiency of HR and GE results and to what degree vulnerable groups are referenced within the evaluation reports. The chapter will conclude with whether the programmatic results capture the quality and the presence of a twin-track approach. The twin-track approach "refers to the strategy to include both gender-specific focus and gender-mainstreaming efforts to ensure that gender equality is integrated."⁴⁴

4.2. Overall Assessment of the UNDAF Contribution to Human Rights and Gender Equality Results

This section examines the extent to which HR and GE programmatic results are integrated in all 50 evaluation reports.

The first component of this meta-synthesis concluded that only four (8 percent) of the 50 reports under assessment were deemed to be gender responsive. Another 27 (54 percent) reports approached requirements. However, as illustrated in chapter 3, these reports did not meet all the criteria to warrant them being considered gender responsive. **The examination of the programmatic results of each report found that the majority of the evaluators did not assess all of the results through a GE and/or HR lens, resulting in a sporadic integration of these perspectives into the evaluation reports.**

With this in mind and as a result of the varying styles, lengths, contexts and approaches to reporting on results of each of the reports, it was decided to focus on four key areas to ascertain the extent to which they had each integrated HR and GE. Four focus areas were selected: governance, economic empowerment, environment/climate change and poverty.

These four areas were chosen as these were the most prevalent among the 194 strategic priorities of the UNDAFs under evaluation and their results could be loosely aligned to the SDGs. In total, 94 strategic priorities⁴⁵ were selected and the subsequent sections were assessed to ascertain the extent to which they referenced GE and/or HR under these four areas. Therefore, the subsequent scores do not relate to the strategic priority as a whole, but only where GE and/or HR is referenced.

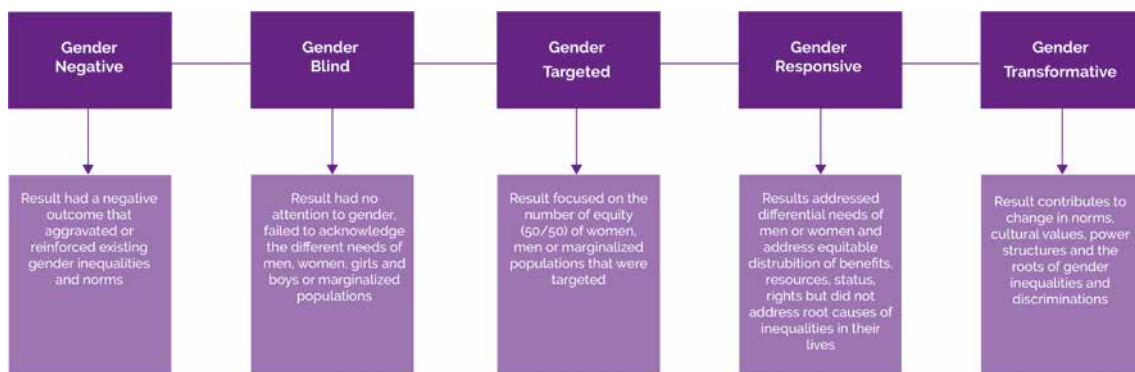
44 Definition taken from United Nations Development Group – Europe and Central Asia, *Desk Review of 18 UNDAFs in Europe and Central Asia, 2018*.

45 The meta-synthesis chose 94 different strategic outcomes – all of these contained at least one element of the four themes selected. Annex VIII illustrates all 194 strategic areas, with the ones being selected highlighted in bold. It should be noted that some of the themes (governance, environment, among others) which were not so obvious in the title of the strategic areas may have been included under other strategic areas. However, this is not included in the assessment.

Each reference under one of these four focus areas was recorded in a database and as reiterated in the methodology, two key analytical frameworks were used to examine the HR and GE-responsiveness of the results of the reports. These included the Gender Results Effectiveness Scale (GRES) and the Gender@Work Framework.

The GRES presents five categories of gender outcomes (Figure 28) which categorize the level of effectiveness or the quality of gender reference or results. The five classifications denote the type of gender change advancement and its effectiveness in transforming gender results.

Figure 28. Gender Results Effectiveness Scale



In order to complement the data and identify the actual change advanced, the Gender@Work framework was also utilized. This framework classifies the type of intervention enabling change and allows the meta-synthesis to assess the extent to which a transformative change had taken place, and whether this was achieved at an individual, formal, informal or systemic level (see under section 4.2.4. for further discussion.)

4.2.1. The Contribution of UNDAF on Human Rights and Gender Equality

Finding Fifteen:

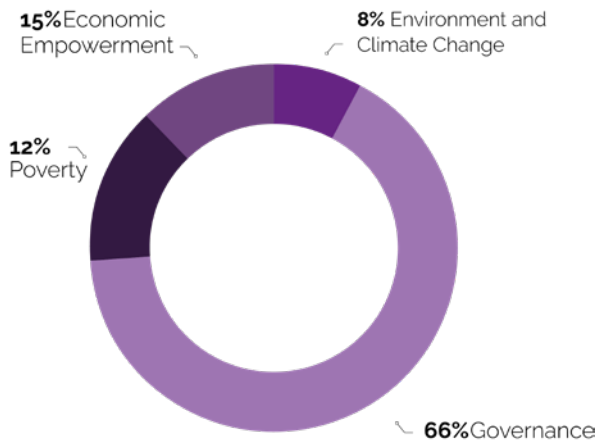
In total, the assessment of the four focus areas (governance, economic empowerment, poverty, the environment and climate change) revealed

443 references⁴⁶ to HR and GE. A total of 334 (77 percent) of these references were found to be gender responsive and therefore had addressed the differential needs of men and women, as well as the equitable distribution of benefits, resources and rights. A further 67 references (15 percent) were deemed gender targeted, while 32 (7 percent) were gender blind.

Only 2 percent of the references were considered gender transformative, which were found solely under governance and economic empowerment

⁴⁶ The meta-synthesis examined the section under effectiveness and scanned the pre-selected sections on the strategic priorities. Each strategic priority under examination was assessed for any reference to gender equality and/or human rights and was thus recorded in a database using the Gender at Work framework. Each reference was then assigned a score under the GRES scale depending on the content of the reference. It should be noted that majority of the references fell under the individual results recorded for the strategic priority. Nonetheless, some reports did not outline the results in an ordered manner, and it was therefore decided to examine each section rather than individual results.

Figure 29. Percentage of gender references per strategic priority



The assessment of the extent to which HR and GE were referenced proved difficult, as the evaluation reports did not follow a standard format. This meant that while some evaluation reports discussed outputs under each strategic priority, others did not refer to specific outputs or outcomes, instead discussing overall results in a generalized manner. To this end, the meta-synthesis included references to HR and GE, rather than examining individual results of the outputs.

The majority of the HR and GE references were found under governance, which accounted for 296 (66 percent) of the results. A further 64 references (14 percent) were found under economic empowerment. Poverty had 47 results (12 percent), while the environment and climate change had only 36 (8 percent). It should be noted that references to HR and GE under other strategic priorities were not included in this particular assessment.

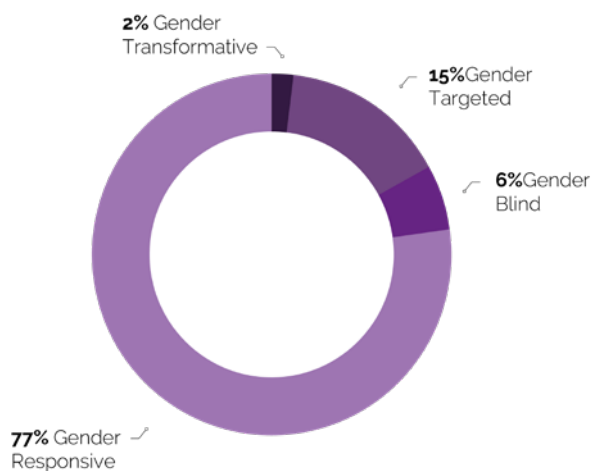
Within the governance area, 84 percent of the references were gender responsive (as opposed to gender targeted or gender transformative). The type of results referenced included the establishment of quota systems to promote women's participation in parliament/Government, the passing of gender-sensitive laws and support to ensuring that governments and civil society organizations reported on the progress of HR and GE to HR bodies. Under poverty, 68 percent of the references were gender responsive, the majority of which centered on access to resources and opportunities. In the area of economic empowerment, 69 percent of

the references were gender responsive, supporting women's empowerment in the workplace, as well as specific skills training to help women gain access to the job market. Within the area of the environment and climate change, 53 percent of the references were deemed gender responsive. Referenced results included the ratification of gender-sensitive climate laws and access to resources on gender-sensitive interventions. This area also had the highest number of results, which impacted on informal cultural norms and exclusionary practices.

As regards the area of economic empowerment, 22 percent were gender-targeted results, which focused on women participating in or benefiting from a certain intervention. Just over a quarter of the results found within the poverty area (26 percent) were deemed to be gender targeted, while governance only had 12 percent of its results classed as gender targeted. The environment and climate change had the lowest number of results (8 percent) classified as gender targeted.

However, the environment and climate change had the highest number of gender-blind results with 39 percent, followed by poverty (6 percent), economic empowerment (6 percent) and lastly, governance (2 percent).

Figure 30. Gender responsiveness of all results per area



Overall Proportion of Gender Outcomes by Areas

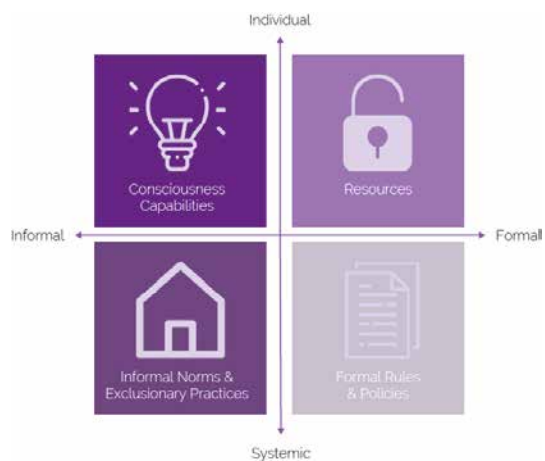
In total, nine results (2 percent) were considered gender transformative. These were found under the focus areas of governance and economic empowerment. One such example occurred in Latin America, where there was a significant change in the lives of rural women, who were able to increase their skills to make decisions about their use of income, while venturing into productive activities. The women also had access to savings banks, thus enabling them to make investments for their children’s education, manage their credit and increase the productivity of their activities.

It should be noted that none of the references were deemed gender negative, though two results from Cape Verde and Kenya were specifically cited as not meeting the targets established in the results matrix. Thus, while the foreseen results would have been classed as gender responsive and gender targeted respectively, the actions resulted in no improvement, with the baselines remaining the same. Although these were the only two cited cases in which targets were unmet, several reports alluded to the fact that they were unable to measure the impact on women’s lives due to a lack of relevant indicators or insufficient data. It is therefore unclear how many of the gender programmatic results may have not met their targets.

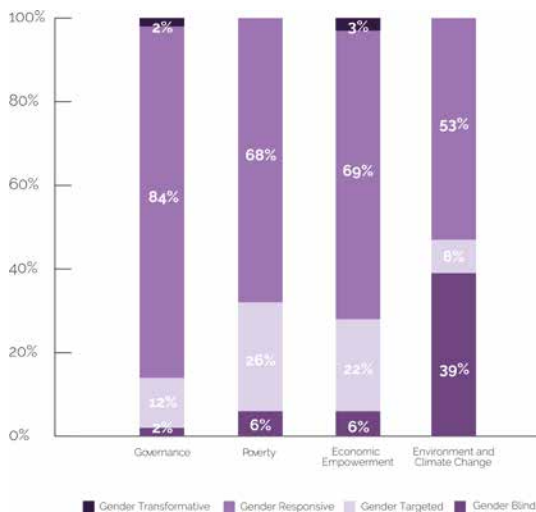
4.2.2 .Key Drivers of Change on Human Rights and Gender Equality

In order to determine what the key drivers of change are, the meta-synthesis uses the Gender at Work framework, which looks at the interlinkages between individual/systemic changes and informal/formal changes. The framework is said to “highlight the interrelationship between equality, organizational change and institutions or “rules of the game” held in place by power dynamics within communities.”⁴⁷ It helps to “identify and connect internal processes to understand and strategize for change across organizational dynamics and broader systems.”⁴⁸

Figure 31. Gender at Work framework



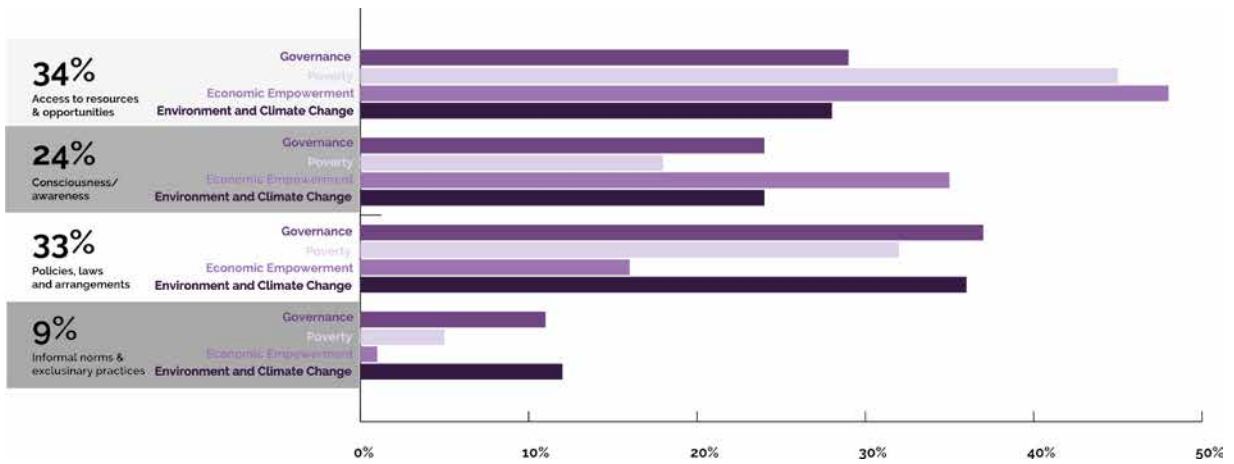
The meta-synthesis utilized the framework by classifying each HR and/or GE-related reference found in the four identified focus areas under one of the four quadrants from the Gender at Work framework. These were subsequently added to a database which further categorized each quadrant under a specific area of work in order to determine the type of intervention used and the actual change recorded.



Overall Proportion of Gender Outcomes in each Result Area

47 Gender at Work, <https://genderatwork.org/analytical-framework/>.
 48 Aruna Rao and others, *Gender at Work: Theory and Practice for 21st Century Organizations* (Oxon and New York, Routledge, 2016).

Figure 32. Proportion of overall gender changes generated



Finding Sixteen:

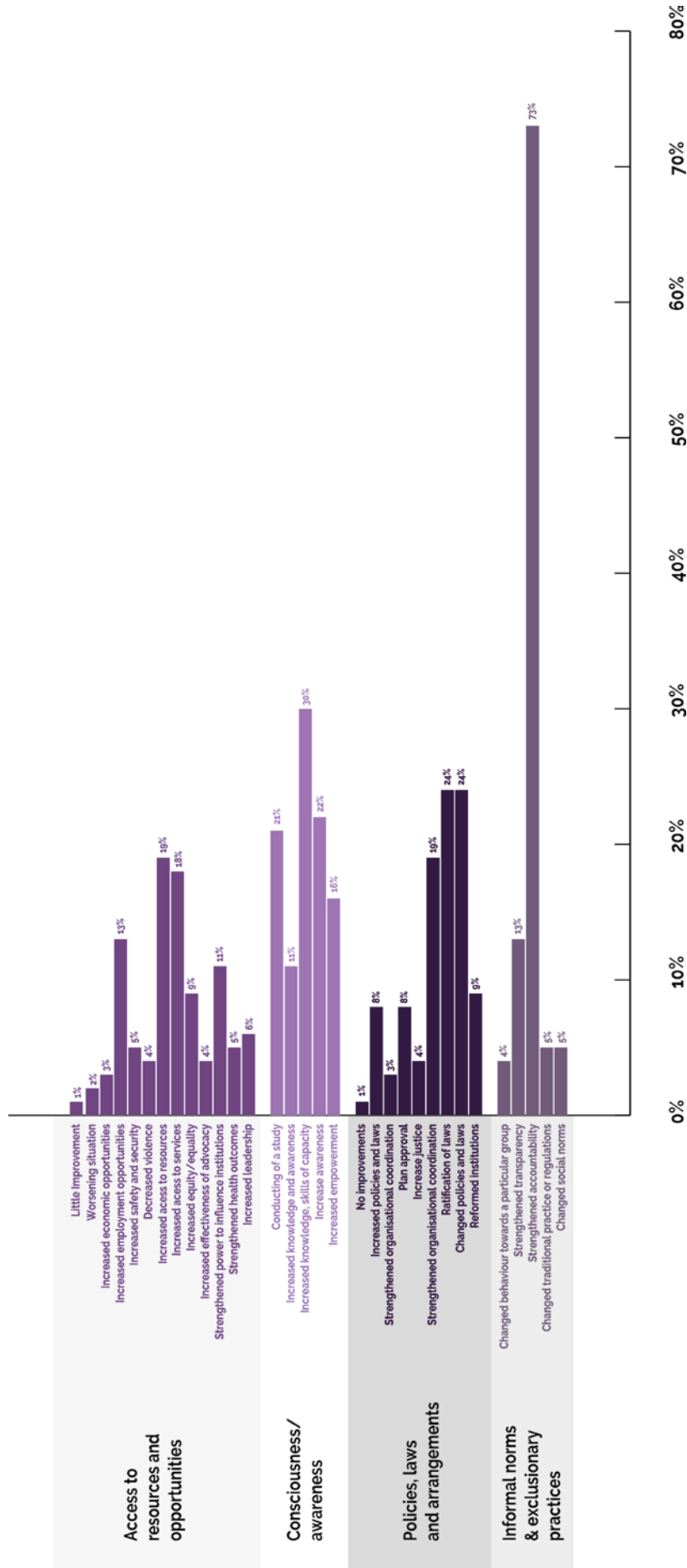
Under the four identified areas of governance, poverty, economic empowerment and the environment and climate change, the results of the assessment concluded that UNDAF contributed equally towards improving gender-sensitive policies and laws (33 percent) and enabling access to resources and opportunities (34 percent) for women and other groups.

Policies, laws and arrangements and access to resources and opportunities share an almost equal percentage in terms of interventions under the four

focus areas in the reports. Almost a quarter of the results indicate that UNDAF supported HR and GE through consciousness and awareness-raising, though only 9 percent of the results contributed to shifts in social norms and changes to exclusionary practices with regards to HR and/or GE.

The findings demonstrate variations between the four quadrants under each of the individual focus areas.

Figure 33. Overview of gender changes generated and gender outcomes by areas



Under policies and laws, the ratification of laws or changes to policies and laws were the most frequent intervention. These normally ensured that laws were more gender sensitive or promoted the participation of women by introducing quota systems. The interventions also contributed to the strengthening of the institutions, ensuring a more gender-sensitive approach in their behaviour, attitude and policies.

The findings revealed that under the area of access to resources and opportunities, the interventions predominantly tended to result in increased access for women and other groups to resources and services, though there were also several interventions that contributed to strengthening the power to influence institutions.

The findings revealed that under consciousness and awareness, the most frequent intervention was enhancing capacity (30 percent), which typically involved training or workshops on a particular subject. Awareness-raising campaigns and studies were also used quite often (21 percent). These studies had the potential to influence policies, though many were also used to influence and the raise awareness of different parties on a particular subject.

Only 9 percent of the interventions caused shifts in social norms and/or shifts in exclusionary practices. Of this 9 percent, an overwhelming 73 percent of the references were deemed to have contributed to strengthening the accountability of institutions.

Finding Seventeen:

The results under the areas of the environment and governance were more evenly spread across the different levels/spheres of change within the Gender at Work framework than the other identified focus areas. In order to ensure transformative change, it is essential to shift attitudinal behaviours and mindsets, which requires a change in internal structures or exclusionary practices, something that is rarely mentioned.

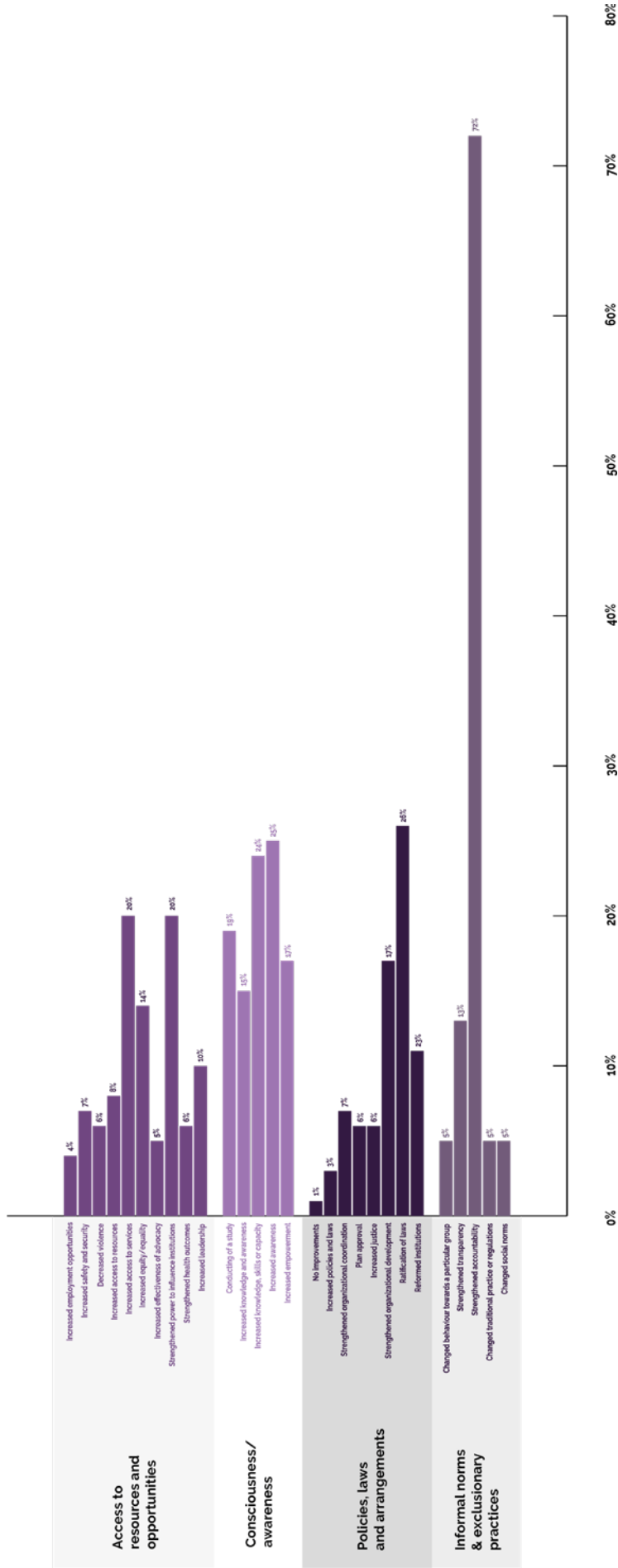
It was uncommon for a singular result or one outcome/result from the same country to include all four types of change. Results tended to focus on a particular area of change, though there were some that indicated several different types of change, with a limited number demonstrating more than one change from a single quadrant. While only nine (2 percent) of results were deemed gender transformative, 77 percent of the results were found to be gender responsive.

Figure 34. Number of results and the changes according to the spheres of work under the Gender at Work framework

No. of results with changes	Governance	Poverty	Economic empowerment	Environment and climate change
One change	242	42	49	22
Two changes	33	5	8	1
Three changes	7	1	2	0
Four changes	0	0	0	0

4.2.3. Human Rights and Gender Equality Results Under the Four Focus Areas

Figure 35. Proportion of human rights and gender equality outcomes and changes generated in the area of governance



Governance

Within the area of governance, a total of 296 references to HR and GE were made. Of the four focus areas, governance had the highest number of gender-responsive references (84 percent).

As Figure 35 shows, governance was also the area that had the most amount of changes to policy and laws (36 percent).

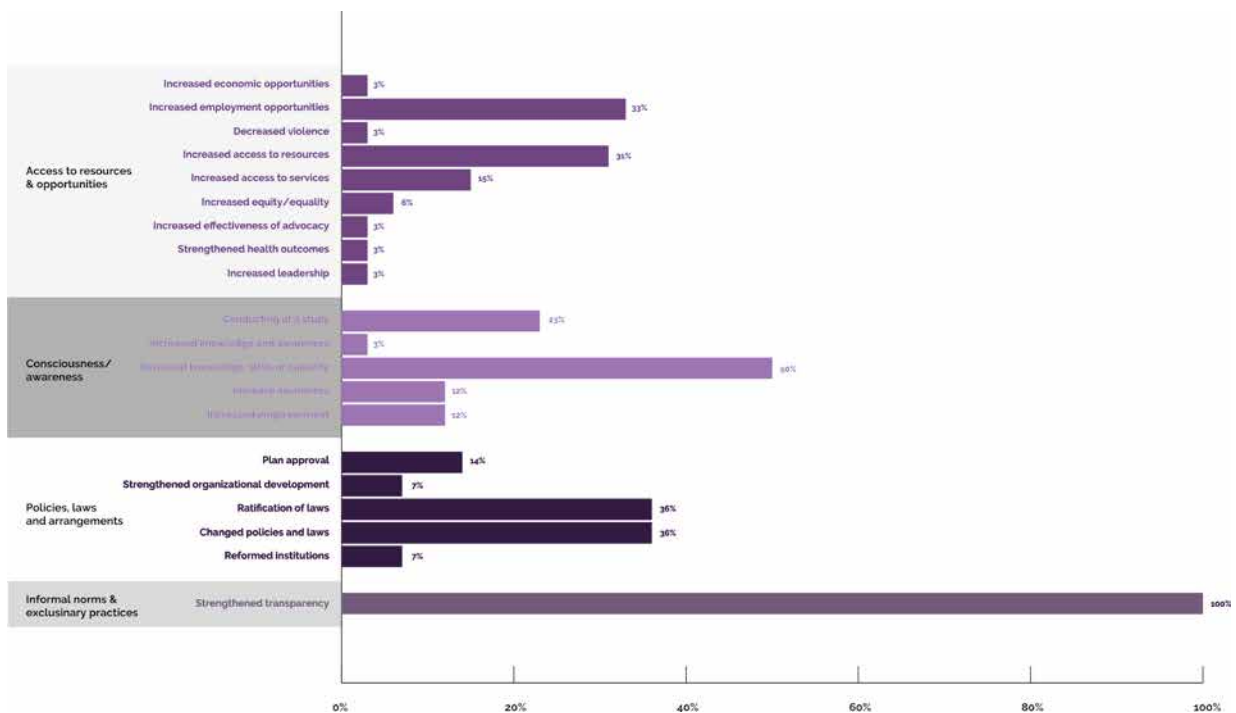
After the environment and climate change, it also had the highest percentage of results which impacted on social norms and internal organizational structures. These were almost always a result of internal structural changes as well as changes in policy and/or law to make them more sensitive to HR and GE, which increased the accountability and transparency of institutions.

Approximately 106 results (36 percent) of the references were aligned with justice. These included the introduction of changes to laws and policies to promote greater gender sensitivity, mostly through improving the legal framework to tackle violence against women.

Other prominent changes included the ratification of HR laws or support for the preparation of HR reports to the various HR bodies.

The area of elections also featured very prominently, with quota systems for women and support for female candidates under political party quotas introduced. The established quotas were generally met, though there was often no indication of how the increased number of women in parliament had actually impacted on ensuring that there are more gender-sensitive laws.

Figure 36. Proportion of gender outcomes and changes generated in the area of economic empowerment



On few occasions, the evaluation reports suggested that indicators were not met and that the number of women being elected had in fact decreased. This was attributed to the fact that the approach to the intervention mainly comprised changes to laws or policies, rather than supporting the capacity-building of female candidates or raising awareness on the importance of having female candidates.

In total, 40 of the 296 changes included more than one change assessed according to the Gender at Work framework. The results revealed that a holistic approach to the perceived discrimination was very rarely taken or at least was not reported upon.

Economic empowerment

This particular focus area accounts for 14 percent of the HR and GE references across the four focus areas. A total of 69 percent of the references were deemed gender responsive, with 22 percent deemed gender targeted. Two of the results were classified as gender transformative, accounting for 3 percent of the overall economic empowerment results. As Figure 36 shows, almost half the changes were attributed to access to resources and opportunities, with outcomes usually resulting in increased employment opportunities and resources.

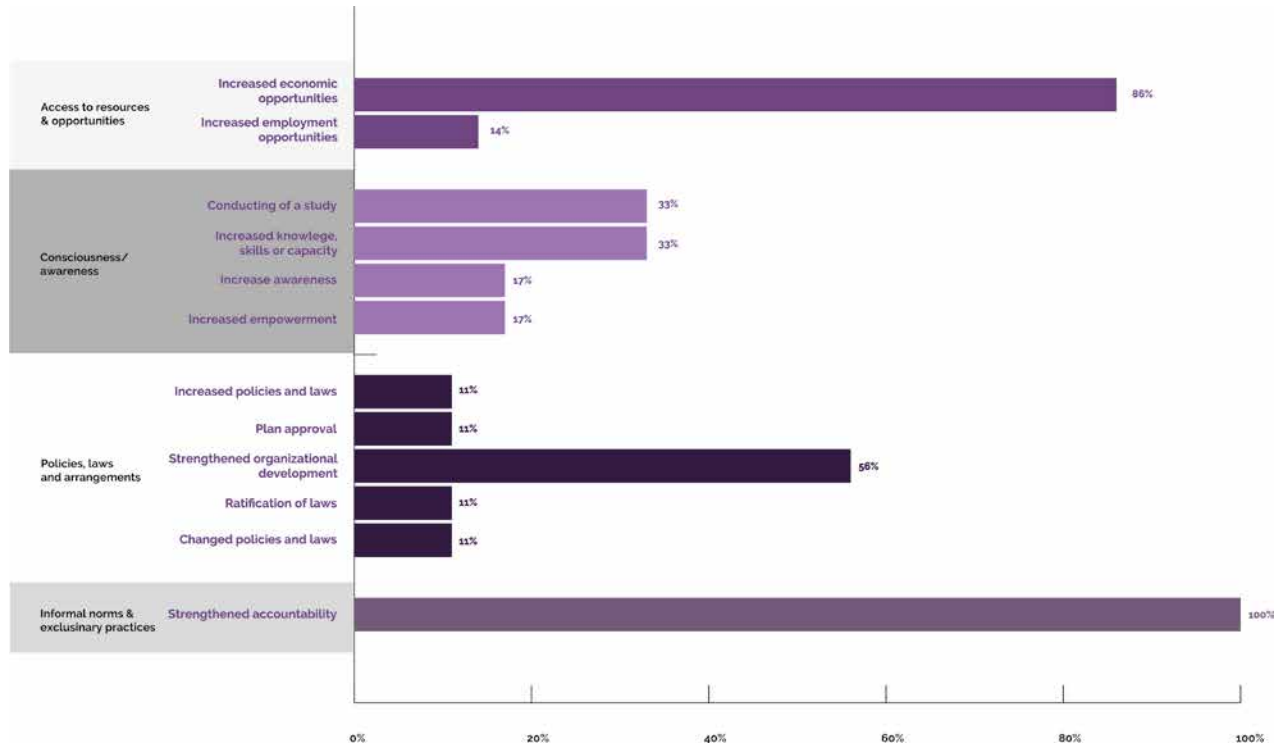
Economic empowerment interventions largely involved women and other vulnerable groups, especially youth, refugees and the rural poor. The beneficiaries of the interventions regularly gained employment and new opportunities were open to them.

Nonetheless, the evaluation report of the UNDAF in Cape Verde described two results whose targets fell below the original baseline indicator, meaning no real results were achieved.

Due to an overall rise in unemployment levels, youth and women's access to employment was impacted upon. Youth were therefore no longer able to access employment and women's entrepreneurship was allegedly reduced due to urbanization and the little remuneration that starting a business had to offer.

In general, many of the references regarding economic empowerment integrated a HR and GE perspective, ensuring that interventions went beyond targeting women and/or a particular group. The interventions to enable access to resources and opportunities were strategic and tended to consider the needs of women and other groups. However, the evaluation reports did not always emphasize what type of impact the inclusion of these groups had and to what extent their inclusion in certain activities had on the lives of the beneficiaries. Although access to resources and opportunities and consciousness and awareness were quite predominant under this focus area, changes in policies and laws were only 16 percent. This meant that while more opportunities and training were being offered, more impact was not possible, as the changes in policies and laws could also lead to change within the other two areas and, more importantly, impact the social norms and cultural attitudes which sometimes act as a barrier for women to access the work place.

Figure 37. Proportion of gender outcomes and changes generated in the area of environment and climate change



The environment and climate change

Results referring to the environment and climate change accounted for only 8 percent of the references across the four areas. Of the references in the programmatic results, 53 percent were found to be gender responsive, though 39 percent were found to be gender blind. No references were deemed gender transformative.

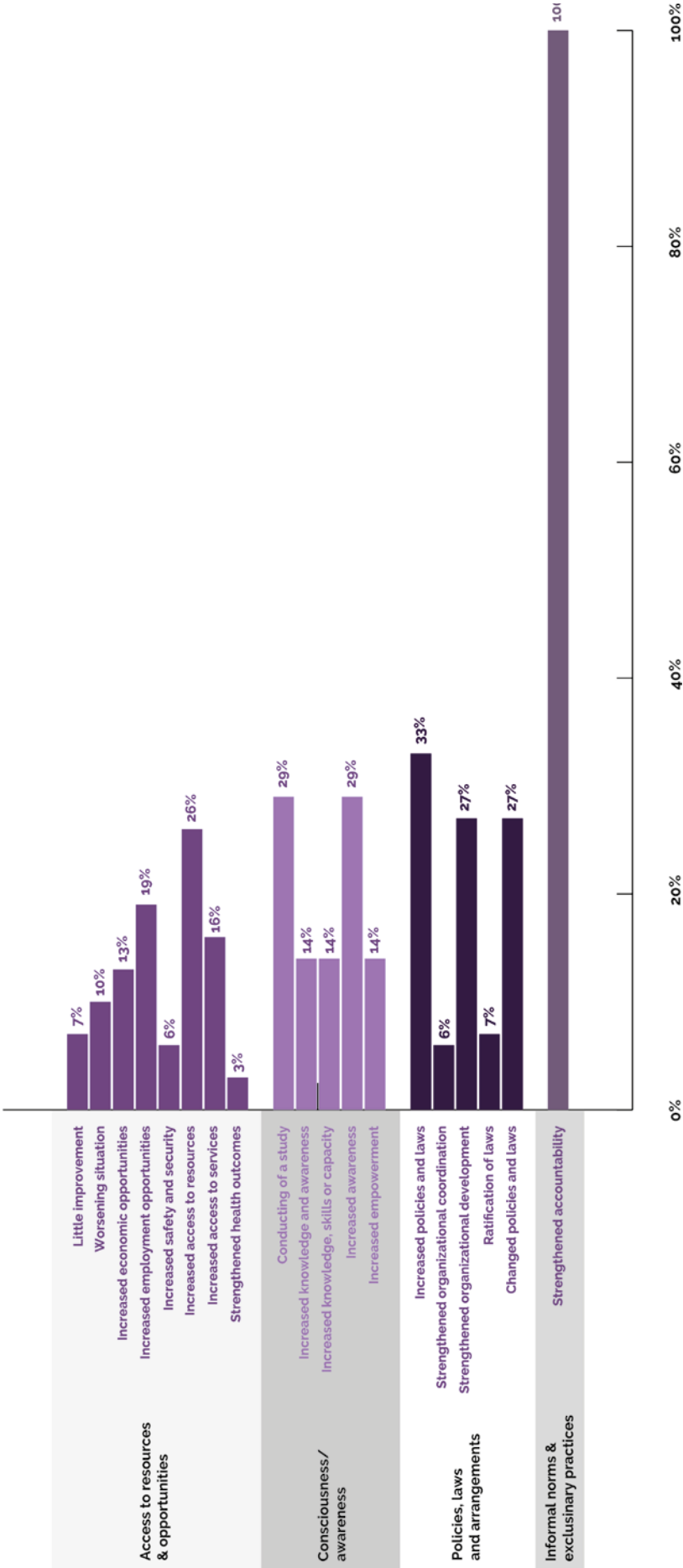
Under the light assessment of the UNDAF frameworks, the strategic priorities that addressed climate change were not found to be gender responsive at all, though several outcomes under these strategic priorities did reference gender and included gender-sensitive indicators.

However, this was not captured in the evaluations, and gains with regard to gender and climate change were rarely mentioned, despite the existence of several gender-sensitive indicators in this focus area.

In 2016, United Nations interventions led to the incorporation of a gender-sensitive approach in the implementation of the Paris Agreement on Climate Change, with the full participation of all stakeholders, especially women.

C.A. Seamus, A. Onwumele and S. Smah,
End of Cycle Evaluation of UN Development Assistance Framework (UNDAF) III, 2014–2017. Final Report. (United Nations Nigeria, 2018).

Figure 38. Proportion of gender outcomes and changes generated in the area of poverty



Nonetheless, Figure 37 indicates that there was a more equal distribution of the types of activities across three of the focus areas of the environment and climate change than in the areas of governance, poverty and economic empowerment. As the figure shows, access to resources and opportunities accounted for the largest proportion. The biggest change, however, was found under laws, policies and regulations, which were seen to have strengthened the organizational development of several key institutions which reportedly became more sensitive to the needs of women and vulnerable groups. Although none of the results were deemed gender transformative, they contributed to some changes in attitudes and enhanced the accountability of the community and local institutions.

Poverty

Poverty accounted for 8 percent of the references in the results across the four areas. As Figure 38 shows, most of the interventions resulted in improved access to resources and opportunities, improving women's and other vulnerable groups' access to the job market or health services.

Around 27 percent of interventions contributed to changes in policies and laws related to women and other groups, which inevitably improved their quality of life. Only 13 percent of the interventions involved awareness-raising, while even fewer (4 percent) ensured that changes were seen in social norms or changes to exclusionary practices.

4.3. The Extent to Which UNDAF has Addressed the Leave No One Behind Principle

FINDING EIGHTEEN: The meta-synthesis revealed that 34 of the reports (68 percent) did not explicitly refer to vulnerable groups, however of the 443 references to HR and GE, 43 percent referenced at least one vulnerable or marginalized group. The majority of these references (144) were found under governance.

As stated at the beginning of this meta-synthesis, the majority of the UNDAFs under examination were developed prior to the SDGs and the development of the leave no one behind principle. However, while the vast majority (49) of the reports did not refer to the principle, 16 evaluation reports did discuss vulnerable groups and other groups which are discriminated against, though this was rarely in-depth, with only two reports referring to vulnerable groups on more than three occasions throughout the entire report.

Figure 39. Number of references to vulnerable groups in the evaluation reports

No. of reports which reference vulnerable groups	
One reference	5
Two to three references	9
Four or more references	2
No references	34

The definition of what constituted a vulnerable group differed according to the UNDAF and country context. Seven reports (30 percent) from Africa referenced vulnerable groups, with two reports (50 percent) from Arab States mentioning such groups.

The Commonwealth of Independent States (CIS) region mentioned vulnerable groups in two out of eight reports (20 percent), which was also the case for the Asia and Pacific region. Latin America referenced vulnerable groups in three out of nine reports. While women, especially victims of domestic violence, rural women and poor women, were highlighted in most of these results in all of the regions, persons with disabilities, refugees and LGBTQ+⁴⁹ groups were rarely mentioned.

Indicators referring to vulnerable groups were not mentioned anywhere in the reports and data were not disaggregated by any particular vulnerable group. Despite this, 18 of the 50 UNDAFs did cite at least one indicator for vulnerable groups.⁵⁰

49
50

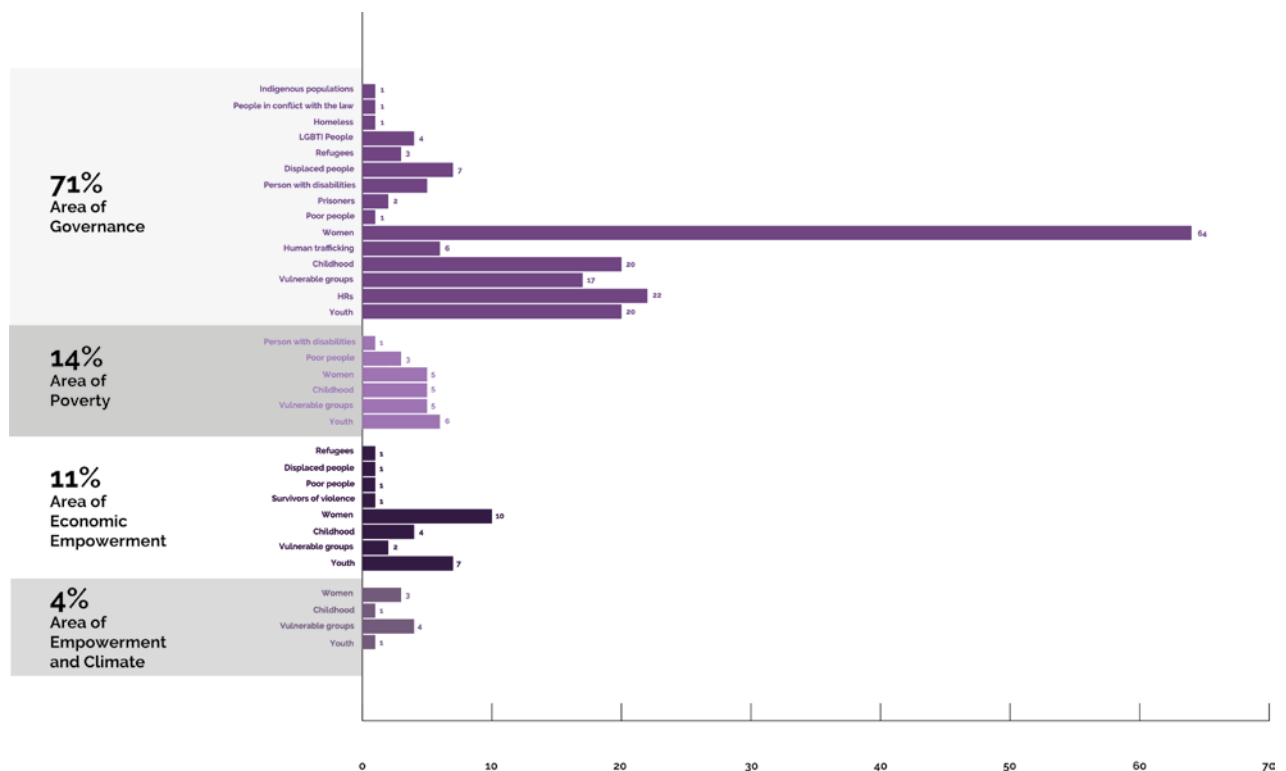
These groups were only mentioned in two of the reports from the CIS region.
Five UNDAFs had either four or more indicators on vulnerable groups, while 12 reports had between one and three indicators.

Figure 40. Number of UNDAF frameworks which reference indicators for vulnerable groups

No. of UNDAF frameworks which reference vulnerable groups in indicators	
One indicator	2
Two indicators	5
Three indicators	5
Four indicators	1
Five indicators	1
Six indicators	1
Seven indicators	1
Eight indicators	1
Nine indicators	1
TOTAL reports	18

As Figure 41 shows, most references to vulnerable groups fall under governance, accounting for 71 percent of the overall results identified. Poverty made up 14 percent, followed by economic empowerment with 11 percent. The environment and climate change accounted for just 4 percent. As noted under governance, the range of groups included is broader than any of the other areas. However, under governance and economic empowerment, it is still women who benefit the most from the interventions, rather than other vulnerable groups.

Figure 41. Proportion of outcomes targeting vulnerable group by area



As previously mentioned, none of the evaluation reports disaggregated data beyond gender, influencing the extent to which many evaluators were able to gauge how UNDAFs had assessed the rights and needs of vulnerable groups.

Furthermore, as was discussed under finding twelve, evaluators rarely consulted with the key rights holders, which included vulnerable groups, largely due to the limited time that the evaluators had in the field. This was further compounded by the fact that meeting a few individuals would not necessarily be representative of the whole group. The possibility to hold focus groups was of course an option, however, the evaluators consulted stated that they had not thought to convene such groups. This inevitably affected how much information could be derived from the consultation through the project documents and interviews with key stakeholders alone.

4.4. Extent to Which Collaborations and Inter-agency Cooperation, Including Joint Programmes and Joint Gender Programmes Contribute to Greater Efficiency on Human Rights and Gender Results

Finding Nineteen:

Collaboration and inter-agency cooperation were discussed in 90 percent of the reports. Countries that had adopted the delivering as one (DAO) approach tended to score higher on the UN-SWAP EPI. However, almost all reports still highlighted weaknesses in terms of inter-agency cooperation, as well as a fragmented approach towards implementation. Most reports did not emphasize as to whether the fragmented approach especially affected HR and/or GE and recommendations to

improve collaboration and inter-agency cooperation were often general and lacking specific focus on HR and/or GE results.

In 2006, the Secretary-General established the High-level Panel on United Nations System-wide Coherence in the areas of development, humanitarian assistance and the environment,⁵¹ which finalized its report in November 2006. One of the key recommendations of the Panel was that the United Nations system should adopt the DAO approach at the country level, with one leader, one programme, one budget and, where appropriate, one office.

The objective of DAO is “to enhance the efficiency, effectiveness and relevance of the United Nations to better assist the Government in achieving its development results.”⁵²

At the end of 2006, eight countries⁵³ began to pilot this approach and in 2019, 57 United Nations offices were classed as DAO.

In total, 30 of the 50 countries under assessment operated as DAO offices. When looking at whether DAO offices help promote gender responsiveness, the impact seems to be clear. **All four reports which met requirements were DAO offices and 18 of the 27 countries which approached requirements were also DAO.** There were eight DAO countries that missed requirements, but five of these reports were deemed unsatisfactory or moderately unsatisfactory.

While many UNCTs were praised for their efforts, especially those operating as DAO countries, 90 percent (45/50) of the evaluations recommended more collaboration between agencies, with reports citing several weaknesses linked to inter-agency collaboration and cooperation.

51 United Nations, “Delivering As One”. Report of the High-level Panel on System-wide Coherence”, 2006. Available at <https://www.un.org/en/events/pastevents/coherence.shtml> (accessed 18 November 2019).

52 United Nations Ghana, Delivering as One in Ghana. <https://ghana.un.org/16048-delivering-one-ghana>.

53 Albania (<http://www.un.org.al/subindex.php?page=details&id=3&mnu=7>), Cape Verde (<http://www.un.cv/en/dao.php#1>), Mozambique (<http://mz.one.un.org/>), Pakistan (<http://unportal.un.org.pk/sites/unpakistan/pages/default.aspx>), Rwanda (http://rw.one.un.org/RWONEUN/index.php?option=com_content&view=article&id=12&Itemid=32), Tanzania (<http://www.untanzania.org/>), Uruguay and Viet Nam (<http://www.un.org.vn/en/the-un-in-viet-nam-mainmenu-37.html>). See <https://undg.org/standard-operating-procedures-for-delivering-as-one/delivering-as-one-countries/>.

This not only affected the assessment of HR and GE results, but results in general.

The key weaknesses referenced were weak monitoring and evaluation (48/50 reports) follow-up and failure to collect adequate data in order to measure the results. This made it very difficult for the evaluators to measure the success of the interventions.

The overall correlation that these weaknesses had on HR and GE results were not always evident, although, as previously stated, the inability to measure results (especially due to the lack of gender-sensitive indicators) hindered the evaluators in measuring them through HR and GE lenses.

Furthermore, although around 25 percent of the reports indicated that GE and/or HR working groups had been established, they were underutilized and had very little impact on results and the UNCTs' ability to promote HR and GE, despite such groups being needed. The reports indicated that the weaknesses in these groups was partly the result of high staff rotation. A lack of interest by some agencies, busy schedules and the need to coordinate an agenda with a large number of agencies and ministries simultaneously for meetings were also cited as some of the reasons why these working groups were not always effective. The evaluation reports did not question their existence, but their effectiveness in influencing HR and GE issues, especially at a programmatic level.

4.5. Evaluation Report Recommendations

The meta-synthesis found that the 50 evaluation reports produced 800 recommendations altogether. However, only 30 of these reports put forward a recommendation containing a reference to gender or women's empowerment. In total, 56 recommendations contained such a reference, with 99 recommendations referencing HR⁵⁴ and 36 referencing the SDGs.

Figure 42. Overview of the type of recommendations for gender equality, women's empowerment and human rights

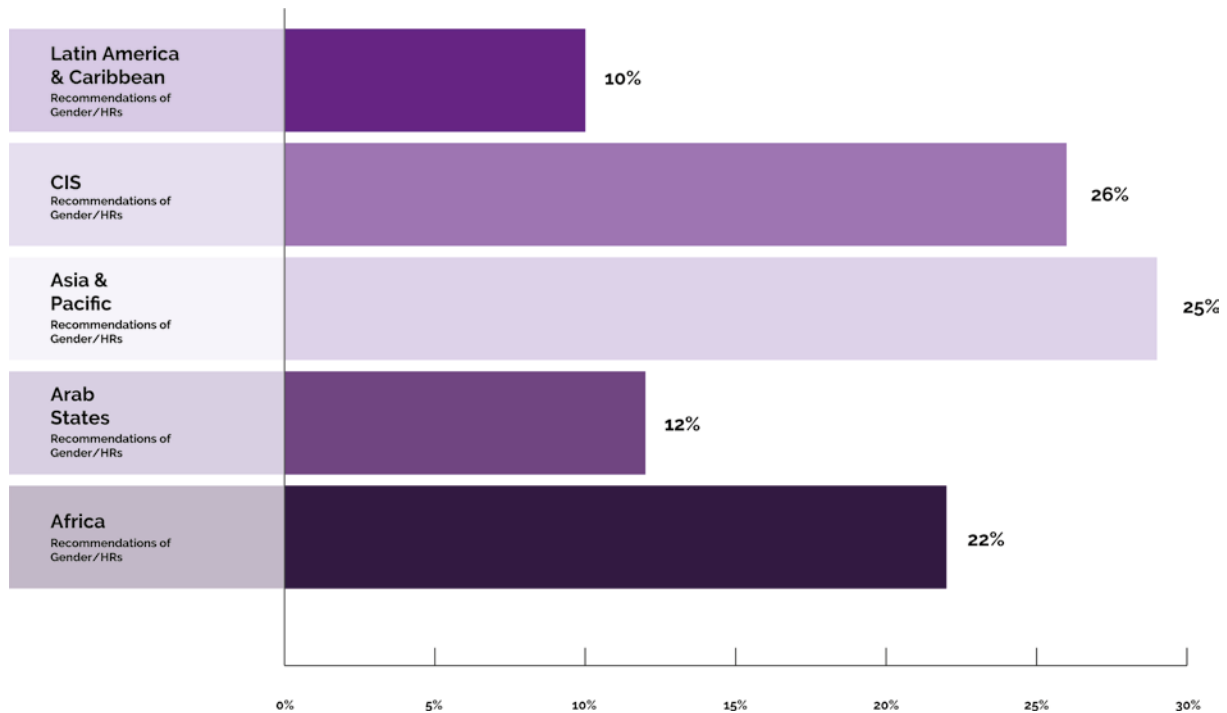
No of Reports with # recommendations	Gender Recommendations	HRs Recommendations	SDGs Recommendations
One Recommendations	16	7	13
Two Recommendations	4	10	7
Three Recommendations	8	6	3
Four Recommendations	1	2	-
Five Recommendations	1	1	-
Seven Recommendations	-	4	-
Thirteen Recommendations	-	1	-

Overall, this meant that the recommendations referencing only gender accounted for 7 percent of the total recommendations, while those referencing both HR and GE totalled approximately 20 percent.

54

These were not necessarily the same reports that had put forward a recommendation referring to gender.

Figure 43. Overall proportion of gender and human rights recommendations



The United Nations Evaluation Group (UNEG) *Guidance on Evaluating Institutional Gender Mainstreaming*⁵⁵ highlights that HR and GE mainstreaming can be implemented in various ways depending on whether it is at the organizational level (institutional versus programmatic) or activity level (for example, policy development versus programme delivery).

In order to analyse the recommendations pertaining to GE, the meta-synthesis divided them into two different categories – programmatic (activity level) and institutional (organizational level), that is to say, policy development versus programme delivery. The recommendations were further categorized utilizing the UN-SWAP 2.0 indicators (see Figure 44) in order to pinpoint the areas that the evaluators had deemed as needing improvement.

Figure 44. UN-SWAP 2.0 indicator areas

UN-SWAP 2.0 Indicator areas

A. Gender-related SDG results

1. Commitment to gender related SDG results
2. Reporting on gender-related results
3. Programmatic results on gender equality and the empowerment of women
4. Evaluation
5. Audit

55

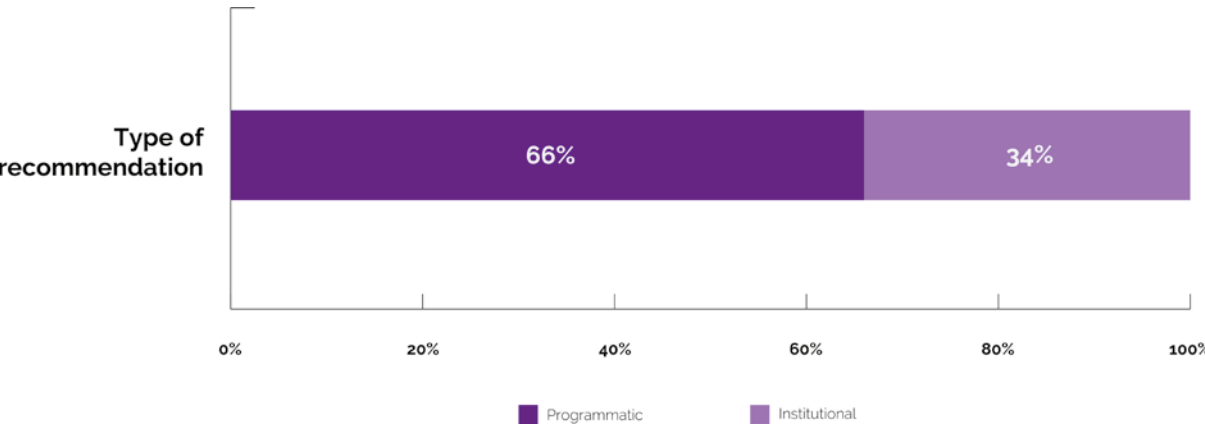
United Nations Evaluation Group (UNEG), *Guidance on Evaluating Institutional Gender Mainstreaming* (2018).

- B. Institutional strengthening to support achievement of results**
1. Policy
 2. Leadership
 3. Gender-responsive performance management
 4. Financial resource tracking
 5. Financial resource allocation
 6. Gender architecture
 7. Equal representation of women
 8. Organizational culture
 9. Capacity assessment
 10. Capacity development
 11. Knowledge and communication
 12. Coherence

Several recommendations addressed the same concerns. However, recommendations for the UNCTs were missing for some weaknesses and problems cited in the reports' findings that were found to be common across the UNDAFs, such as, for example, the quality of the indicators and the lack of adequate monitoring and evaluation.

In total, there were 31 reports which provided a total of 56⁵⁶ gender-related recommendations. As Figure 45 shows, most of the recommendations were programmatic and related mainly to the development and design of the UNDAF, rather than its overall implementation, although there were some evaluation reports whose findings cited that implementation of gender-responsive results were weak, though several reasons have been suggested for this. Some of these reasons relate to the context, security situation in the country and a weakness in the results matrix, each of which prevent the proper monitoring of results. These findings were not always reiterated in the recommendations.

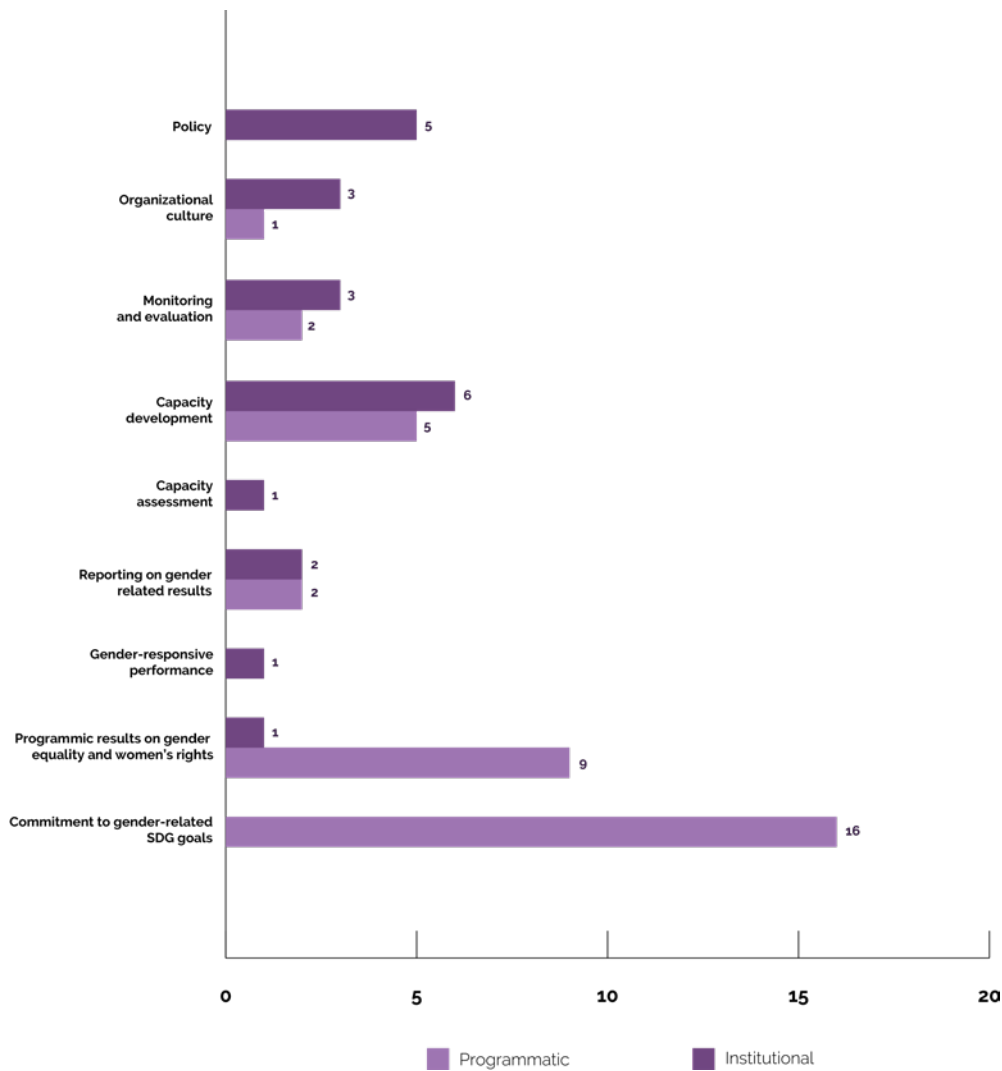
Figure 45. Type of gender-related recommendations



Just 34 percent of the recommendations (Figure 45) were interpreted as institutional, but these were usually in relation to capacity-building, internal policies or leadership.

⁵⁶ Overall there were 52 recommendations, but in order to classify them according to the UN-SWAP, some were divided giving a total of 56.

Figure 46. Overview of the type of gender-related recommendations



The following section provides an overview of the recommendations according to the assigned category. The analysis also refers to how often references are made to these issues in the findings of the reports. As previously mentioned, some of the weaknesses with regards to gender-related results were not deemed as requiring a separate and/or explicit recommendation by evaluators and were only highlighted within the findings.

In total, 56 gender-related recommendations were made, 37 of which were found to be programmatic, with the remaining 19 deemed institutional. The recommendations could be categorized under the following in line with the UN-SWAP 2.0 guidelines:

Commitment to gender-related SDG results

This examines whether a particular result supports gender equality standalone work and mainstreaming of equality policies, programmes and institutions.⁵⁷

Of the 56 recommendations, 16 referred to the importance of ensuring that gender was mainstreamed into other areas, such as education, justice and security. Although the meta-synthesis found that most of the references to HR and GE in the results (74 percent) were gender responsive, the integration of HR and GE perspectives throughout the findings was quite low. HR and GE results were either largely scattered throughout the reports or found within an additional paragraph that assessed HR and GE. To this end, a little over half the reports were ranked as only approaching requirements. HR and GE results were integrated into almost every type of strategic priority in only a handful of reports (mainly those that met requirements), despite some UNDAF results frameworks not being particularly gender responsive.

Incorporate the principles of environmental sustainability in all the priority strategic areas of the UNDAF to add value to development work and advance the gender equality and human rights agenda. Increase the understanding of planners about the links between the environment and poverty, migration, agriculture and regional development, etc. Environmental protection and sustainable use of natural resources is also a matter of gender and human rights.

United Nations Moldova, Final Evaluation of UN-Republic of Moldova Partnership Framework 2013–2017. Final Evaluation Report (vol.1). United Nations Republic of Moldova, 2016.

Reporting on gender-related results

This refers to the importance of how results-based management should address the design of gender-sensitive results and indicators.⁵⁸

Recommendations to improve the mainstreaming of gender across the entire programme in order to ensure that results advance GE were put forward by three different countries and accounted for four of the recommendations (7 percent). Two of the countries identified the problem as relating to the UNDAF itself,⁵⁹ while the other country⁶⁰ put forward two recommendations to recruit consultants with expertise in gender who could participate in the planning of the new UNDAF with expertise in results-based management.

Despite only three reports providing recommendations on this issue, the majority of reports discussed at least one problem within the findings with regards to the indicators of the UNDAF priority areas. Only nine out of 10 evaluation reports cited disaggregated data by gender, the majority of which were found within the background section. However, there were also a number of references (24) highlighting the effectiveness of the indicators.

57 UNEG, *Guidance on Evaluating Institutional Gender Mainstreaming* (2018).
58 Ibid.
59 Angola and Afghanistan.
60 Somalia.

Figure 47. Number of references to disaggregated data in the evaluation reports

No. of references (Rf.) in reports which disaggregate data by gender		
Rf. background	Governance Economic empowerment Security Social Services Results framework	15
Rf. relevance		0
Rf. effectiveness		3
		7
		1
		5
		26
Rf. efficiency		0
Rf. sustainability		0
Rf. impact		0
TOTAL reports		9

The problem of ensuring gender-responsive indicators and reporting on gender-related results was also highlighted by five of the Resident Coordinators consulted. While many UNCTs have gender focal points, they tend to be junior staff who also work on other tasks and do not always have specialized gender training. However, it is worth noting that gender experts often lack certain training, such as monitoring and evaluation for example, with few experienced in designing gender-responsive indicators. **A lack of gender-responsive indicators is definitely an important factor as to why many evaluators felt they were unable to integrate HR and GE to a larger extent in the evaluation reports.**

Policy

This states that United Nations entities should have a current gender equality empowerment policy or plans in place that address gender mainstreaming, GEWE and HR.

Establish indicators within each priority area that measure each cross-cutting issue alongside joint programming to apply a cross-cutting approach.

Saiyed M. Husain, United Nations Development Assistance Framework Afghanistan (2010–2014). Final Report. (United Nations Afghanistan, 2015).

Four of the reports (Gambia, Jordan, Kenya and Ukraine) highlighted the need to develop a UNCT-wide gender strategy in order to articulate a consistent approach for integrating, monitoring and measuring results in the next UNDAF cycle and for integrating the gender programming principle, including through the implementation of the UNCT-SWAP Gender Equality Scorecard.⁶¹

⁶¹ "20. Gender equality and women's empowerment are integral to realizing the 2030 Agenda and all the SDGs. To integrate a focus on these issues throughout the Cooperation Framework, UN development entities should put gender equality at the heart of programming, driving the active and meaningful participation of both women and men, and consistently empowering women and girls, in line with the minimum requirements agreed upon by the United Nations Sustainable Development Group (UNSDG) in the UNCT System-wide Action Plan (SWAP) Gender Equality Scorecard." United Nations Sustainable Development Group (UNSDG), United Nations Sustainable Development Cooperation Framework. Internal Guidance (2019). Available at https://undg.org/wp-content/uploads/2019/06/UN-Cooperation-Framework-Internal-Guidance-25_June-2019.pdf

While these were the only four reports that highlighted this recommendation, several other reports cited the need for better cooperation between the United Nations agencies to integrate gender within the findings (under efficiency).

Leadership

This emphasizes the importance for senior leaders to have a clear vision in promoting GEWE.⁶²

This recommendation was only highlighted in two reports.⁶³ The remaining reports did not highlight leadership as an issue, in fact, several instead reported that the Resident Coordinators had "good leadership," which had contributed to the success of the UNDAF. Nonetheless, this was not necessarily in relation to the extent to which they supported HR and GE.

The UNS [United Nations system] should continue its advocacy work on gender violence to mobilize more actors and integrate this issue into several sectors. One should also ensure that this issue is addressed in the next PND [National Development] Action Plan through strong political leadership and a multi-year action programme, mobilizing all the vital forces of the country in all productive and social areas.

Moncef Kouidhi, Alban Ahoure and Raymond Dogore. UNDAF 2013–2015. Final Assessment. Knowledge, Lessons Learned and Strategic Approaches for 2017–2020. Final Report. (United Nations Côte d'Ivoire, 2015).

The context of the country is an important factor, though at times it is simply too difficult to factor in gender when a government is particularly reticent, especially with regards to human rights and groups such as the LGBTQ+ community.

Furthermore, as previously highlighted, there is sometimes a need to be more cautious with regards to HR and GE, especially with regards to data, as a lack of protection and confidentiality around sensitive issues can endanger individuals. It is not always possible to collect disaggregated and reliable data.

Financial resource allocation

In order to ensure adequate resources for GEWE and gender mainstreaming, United Nations entities must establish a financial benchmark for resource allocation in all of its budgets, including headquarters, regular, core and extrabudgetary resources.⁶⁴

There were only three countries⁶⁵ that referred to the gender allocation budget in the recommendations, although budget and a lack of resources was cited as a general problem in at least a third of the reports. Five reports explicitly cited that resources had been a problem and as a result, the HR and/or GE referenced result had not been achieved.

Establish an allocation budget for activities in the new UNDAF Action Plan (for example, at least 2 percent of programme budgets are assigned for the integration of the normative agenda, including previous analytical work on gender, human rights and other central problems).

Bangladesh

62 UNEG, *Guidance on Evaluating Institutional Gender Mainstreaming* (2018).

63 Côte d'Ivoire and Sri Lanka.

64 UNEG, *Guidance on Evaluating Institutional Gender Mainstreaming* (2018).

65 Bangladesh, Guatemala and Senegal.

Capacity development

*GEWE training should be mandatory for all staff with additional training offered for gender focal points, specialists and managers.*⁶⁶

Capacity development was highlighted in 11 of the reports, with nine recommendations made. Some of these recommendations referred to the actual UNDAF results, such as that women and other groups should undergo capacity-building, and therefore did not fall under the SWAP indicator.

Nonetheless, six other reports recommended that the UNCT should ensure capacity development, mainly to ensure that HR and GE are integrated into the UNDAF, as well as support for staff to develop and design the indicators.

The UNCT should undertake special efforts, including appropriate capacity development of staff, in order to ensure the integration of programming principles, viz., human rights-based approach (HRBA), gender equality, environmental sustainability, results-based management and capacity development in the next UNDAF.

Elballa Hagona, Evaluation Report of Sudan UNDAF (2013–2016) (United Nations Sudan, 2015).

4.6. Key Challenges to Ensuring a Gender Responsive Evaluation

The meta-synthesis has concluded that there are several challenges in meeting requirements for conducting gender-responsive evaluations. Chapter 2 concluded that while some of these challenges lay with the evaluators themselves, especially with regards to ensuring the inclusion of a gender-responsive methodology and scope, other factors also play an important part.

This section aims to capture some of the highlighted challenges to integrating HR and GE found within the evaluation reports.

Finding Twenty:

The evaluation reports cited various references that contributed to the integration of HR and GE in the UNDAF results. However, the overall impact on HR and GE is sometimes difficult to gauge due to poorly designed results frameworks and/or the unavailability of information or data.

In general, it is difficult to ascertain the true extent that the UNDAF impacted on HR and GE and how the results that showed HR and GE integration have led to a transformative change.

The difficulty in extracting this information from the evaluation reports lends itself to the fact that the reports tended to list the interventions without necessarily delving into whether the interventions had been effective in addressing HR and GE and to what extent. This is compounded by the fact that evaluators concluded that it was difficult to gauge the success of certain results, both general results and HR and GE-targeted results, due to the lack of baseline data and adequate indicators.

66

UNEG, *Guidance on Evaluating Institutional Gender Mainstreaming* (2018).

However, this conclusion is impacted by weaknesses in the Results Framework, a lack of clarity in the definition of status markers, and a lack of access to available data on which to accurately assess the status of outcome indicators against established baselines and targets.

Ann Lund and Betty Lovai, Independent Evaluation. UN Development Assistance Framework (2012–2017) (United Nations in Papua New Guinea, 2016).

As the analysis demonstrates, within the area of governance, there was a significant number of references to results that had integrated HR and GE perspectives and that showcased how these results had helped improve the participation of women in parliament or the Government. However, it was not always evident as to what extent the UNDAF itself had contributed to the result or indicator being met. Furthermore, several reports stated that indicators were set too high and were therefore difficult to measure, making it near impossible to determine the direct contribution made by the UNDAF. This was reiterated by some of the evaluators also consulted, especially with regards to results under governance.

[S]ome of the indicators were not sufficiently measurable. The indicators for democratic governance, in particular, were not measurable. It would have been possible to identify alternative indicators, which could have been both qualitative and measurable.

UDA Consulting, Evaluation of the UNDCS for Turkey (2011–2015) – Final Report (2016).

Furthermore, unlike evaluations on individual projects, the UNDAFs are far more fragmented and a theory of change was not always evident. The theory of change is now an important aspect of the new wave of Cooperation Frameworks, which recent guidelines have highlighted. According to one evaluator, the UNDAFs and Cooperation Frameworks are not very clear with regards to defining a HR or gender approach, suggesting that HR and GE are almost an afterthought in some of the result matrices.

To this end, it is paramount that the theory of change takes into account the development of gender-sensitive indicators. As stated in the report of the Secretary-General on mainstreaming a gender perspective into all policies and programmes in the United Nations system, “without supportive surge capacity, particularly during the critical design and evaluation stages, progress will remain stalled.”⁶⁷

The report further noted that many of the UNCTs struggled to demonstrate results owing to, among others, a lack of gender-sensitive monitoring systems. Only five UNCTs met the requirements for the UNDAF monitoring and evaluation indicator to measure progress against GE results.⁶⁸

Finding Twenty-One:

Evaluators and UNCTs face several challenges in ensuring that the UNDAF is successfully evaluated. Nonetheless, the overall ability to ensure the integration of a HR and GE perspective into the evaluations principally depended on the quality of the results matrix and the availability of information, including sex-disaggregated data, to successfully assess the indicators and results.

67 E/2019/54.
68 Ibid.

One of the key elements that was commonly mentioned was the difficulty in obtaining data in general, as well as the fact that UNCTs had often failed to monitor all indicators established in the results framework. The lack of this information affected the ability to conduct an effective evaluation, especially with regards to obtaining information on HR and GE results. This was seen as a particular challenge by all of the survey respondents and those who were interviewed.

The Resident Coordinators consulted also pointed to the difficulty in sometimes keeping track of indicators as well as the weak commitment and capacity of some agencies and especially governments to obtain and maintain accurate databases and figures, particularly those disaggregated by gender.

The shortfall of HR and GE experts in the UNCT and their ability to design an effective results framework⁶⁹ could also have contributed to the fact that many of the results matrices were deemed weak. Another possible factor was that although 20 reports mentioned that civil society had been consulted during the design of the UNDAF, only three alluded to civil societies' involvement thereafter. It is unclear whether this included support to design the Monitoring and Evaluation Framework, especially with regards to the development of the indicators in the results matrix.

To this end, according to some of the evaluators consulted, HR and GE were not articulated often, and the results matrix did not indicate how HR and GE would transverse different outcomes.

While UNCT offices do ensure that gender issues are considered a priority, this is not necessarily the case in practice. The enormous potential of the subject and the real issues that exist deserve more value, which is needed to achieve greater traction.

This was further exacerbated by the fact that another key critique of the UNDAF, as previously mentioned, was that its results framework had far too many indicators to be monitored effectively and in several instances lacked baselines for the indicators. The number of indicators ranged from 12 to 348, making it almost impossible for evaluators to collect all the relevant information in the short time they had to conduct the evaluation. This issue was further emphasized by the fact that several countries failed to produce annual reports and, in some cases, according to one Resident Coordinator consulted, there was not enough demand for information about UNDAF performance. Four of the Resident Coordinators consulted confirmed that it was often difficult to ensure effective monitoring and evaluation of results and that the inclusion of monitoring HR and GE results was almost considered secondary, as there were no effective systems in place to ensure this.

Furthermore, while some governments used gender disaggregated data, some of the Resident Coordinators mentioned that governments were sometimes hesitant to use the same indicators as those utilized by the United Nations. In some instances, this reluctance was due to governments' lack of capacity to collect data for particular indicators, though on some occasions, it related to their resistance to include certain HR indicators.

The results had limited success as governments did not necessarily prioritize HR or GE to the same extent as the United Nations, which was one of the main challenges for gender-focused interventions.

One Resident Coordinator cited that it was sometimes difficult to promote HR in the country context, so their office tended to package HR under the leaving no one behind principle, as the Government was more open to this.

69 Several of the RCs consulted stated that while some offices did have gender experts, they were not necessarily experts in monitoring and evaluation and had very little knowledge on the design of gender-sensitive indicators. Furthermore, most of these gender focal points were either junior staff members or were responsible for a number of tasks, meaning they were unable to assign more time to the development of the UNDAF results matrix.

Adopting such a strategy was not without challenges and meant that some rights were simply not included as they were considered too sensitive and could not be 'packaged' under the leaving no one behind principle. As a result, indicators that measured beyond gender were rarely designed and data on vulnerable groups, especially the LGBTQ+ community, were seldom collected within many UNDAFs.

The extent to which a government pushes back on the rights agenda is thought to be largely dependent on the context of the country. Some Resident Coordinators commented that they had experienced not having the power to defend and advocate for HR for fear of causing too much offence. This has sometimes meant that UNDAFs have not been reflective of a HR and GE approach, simply because the UNDAF needs to be endorsed by the Government.

4.7. Twin Track Approach

Finding Twenty-Two:

The light assessment of the UNDAF results indicated that the application of the twin-track approach was not well balanced. This finding is substantiated by a significant number of reports that found that HR and GE were not sufficiently mainstreamed across UNDAF evaluations.

Reports alluded to weaknesses in the twin-track approach, highlighting that HR and GE were rarely mainstreamed across the UNDAF outcomes.

In Montenegro, for example, the country's report noted that there is no evidence of a HR-based approach in the pillar working groups of the Steering Committee, and that HR issues are mainly addressed at the agency level. As such, the HR and GE working group has a limited mandate, outreach and visibility.

In Cuba, the country's report indicated that issues within the UNDAF lack focus and that gender-related issues are not clearly characterized, despite numerous studies and work carried out by the Interagency Group. The enormous potential of the subject and the real issues that exist deserve more value, which is needed to achieve greater traction. In a similar vein, Honduras noted that while gender issues are considered a priority, this not reflected in practice. In fact, the general perception seems to be that the staff's approach is not properly understood, which is a clear reflection on the UNDAF's design.

The Resident Coordinators consulted all confirmed that they prioritized HR and GE, though many suggested that the UNDAFs and their results were not always taken seriously. There was a tendency by some offices to forget about the UNDAF once it had been developed. Agencies all had their own agenda and there was not always the commitment needed to ensure that the results envisaged in the results framework would come to fruition.

V. Conclusions and Recommendations

Through its findings, this meta-synthesis has established that there are several influential variables that determine the extent to which a UNDAF evaluation carried out between 2015 and 2019⁷⁰ integrated HR and/or GE perspectives.

The meta-synthesis has demonstrated that while efforts were made to integrate HR and GE in most of the reports, it still only led to four of the 50 reports (8 percent) meeting requirements according to the UN-SWAP EPI.

Although this is only a small percentage, an additional 27 reports (54 percent) were classified as approaching requirements, 19 of which (70 percent) were considered as potentially being able to meet the required standards had they included a stronger HR and GE design and methodology. In addition, these 19 reports included recommendations on either HR or GE and in some cases put forward recommendations on both.

Before highlighting the key recommendations of this report, it is important to underline that the meta-synthesis covered UNDAFs rolled out between 2010 and 2016, with only five of these rolled out in 2015 or 2016. The introduction of the SDGs in 2015 and the 2030 Agenda has marked an important shift in the way that UNDAFs and Cooperation Frameworks are designed, which may impact the overall quality and integration of HR and GE in evaluations of such frameworks in the future.

Furthermore, in 2019, General Assembly resolution 72/279 elevated the UNDAF (now renamed UNSDCF)

as “the most important instrument for the planning and implementation of United Nations development activities in each country, in support of the implementation of the 2030 Agenda for Sustainable Development.”⁷¹

This contrasts substantially with the findings of the meta-synthesis, as the desk review and key informant interviews alluded to the perceived weakness of the UNDAFs. It is hoped that the elevated status of Cooperation Frameworks will stop them from being ‘shelved’ or subject to ‘weak’ results monitoring in the future.

To this end, it is anticipated that some of the highlighted weaknesses and challenges facing UNCTs may improve following the restructuring of RCOs. The prominent position that the Cooperation Framework has over the entire programme cycle may also lead to improvements, since this encourages United Nations development entities to place HR and GE at the heart of programming.

However, while some of the aforementioned challenges may easily be addressed, several of the key highlights of this report have already been outlined in existing gender-responsive guidelines. It should be noted that this report concludes that there is no single variable that can effectively contribute to a HR and gender-responsive evaluation. In fact, to conduct high-quality evaluations requires several variables, making it paramount for UNCTs and RCOs to continue building on their professionalism and investing in specialized knowledge, education and training to enhance the quality of HR and GE evaluations.⁷²

⁷⁰ Only one evaluation from 2019 was selected.

⁷¹ A/RES/72/279.

⁷² UNEG, *Evaluation Competency Framework* (New York, 2016).

Evaluations should be conducted with professionalism and integrity. Professionalism should contribute towards the credibility of evaluators, evaluation managers and evaluation heads, as well as the evaluation function. Key aspects include access to knowledge; education and training; adherence to ethics and to these norms and standards; utilization of evaluation competencies; and recognition of knowledge, skills and experience. This should be supported by an enabling environment, institutional structures and adequate resources.

UNEG, Evaluation Competency Framework (2016).

Integration of HR and GE perspectives in UNDAFs

HR and GE perspectives are increasingly being integrated in the UNDAFs and Cooperation Frameworks, which is inevitably facilitating the integration of HR and GE perspectives into the evaluations. Under finding one, the meta-synthesis revealed that UNDAFs classified as highly gender-responsive (i.e. over 60 percent of outcomes detailed in the results framework matrix were deemed gender responsive), tended to be ranked higher on the UN-SWAP EPI and were therefore considered to integrate HR and GE to a higher degree.

Aside from this correlation, the findings of the meta-synthesis concluded that the twin-track approach was not necessarily applied to all results and while there were several outcomes that were HR and GE-responsive, both elements were clearly not mainstreamed across the entire UNDAF. The mainstreaming of HR was also deemed very weak, with relatively few references made to HR in the reports' findings. Furthermore, as the second component of the meta-synthesis revealed, HR and gender-specific results tended to revolve around a limited number of strategic priorities, meaning that

some strategic priorities did not mainstream HR and/or GE at all. This was evident in the findings of the 50 evaluation reports, in which references to HR and GE were largely concentrated in the focus area of governance.

Recommendation One::

Efforts should be made to ensure that HR and GE are sufficiently mainstreamed across Cooperation Frameworks. This includes the need to guarantee more HR and gender-sensitive indicators in order to measure the impact that interventions have had. Efforts should concentrate on ensuring Cooperation Frameworks are balanced, thus enabling HR and GE to be mainstreamed across all strategic priorities, instead of just a few key areas.

UNCTs commitment to mainstreaming HR and GE in the evaluation reports

Finding two, which was determined through the desk review and subsequent consultations with Resident Coordinators and evaluators, concluded that the UNCTs strive to promote HR and gender-responsive evaluations, but could improve several factors to yield more positive results.

The meta-synthesis revealed that despite clear guidelines on how to integrate HR and GE in evaluations, the UNCTs experience several challenges in ensuring that evaluations are sensitive to HR and GE.

Human resources

Adequate human resources are a key element in ensuring a HR and GE approach to evaluation.

UNCTs often lack sufficient gender expertise or a dedicated staff member responsible for overseeing that HR and GE are integrated throughout the evaluation process. Gender focal points tend to be junior staff who are assigned to varying duties, meaning they often lack the capacity and time needed to address HR and GE in the UNDAF and Cooperation Frameworks.

Limited human resources not only hampers the evaluation process, but also impacts on the entire design and implementation process related to the UNDAF Cooperation Framework.

Design of the UNDAFs and Cooperation Frameworks

The desk review revealed that the UNDAFs under evaluation were all relevant to national and international priorities and a wide array of relevant stakeholders had been consulted in their development. Furthermore, there is solid evidence from the desk review that HR and GE have been integrated into the UNDAF frameworks, aligning with the national and international priorities and commitments.

Nonetheless, the desk reviews principally concluded that the design of the results matrix appeared to be weak, which was particularly apparent with regards to HR and GE.

In general terms, the indicators tended to be over ambitious, unrealistic or simply not achievable. To this end, the ability to measure the impact of the UNDAF during an evaluation at both the outcome and output levels was substantially weakened as a result of a poor results framework. This was further accentuated by the lack of information available, as well as governments' and sometimes UNCTs' inability to adequately track results.

Only a small percentage of reports (three reports) referred to a lack of HR and gender-sensitive indicators, although upon examination of the UNDAFs very few HR and gender-sensitive indicators were found to have been included. The fact that evaluators did not report on this is revealing, as it demonstrates a lack of awareness on the importance of ensuring a HR and gender-sensitive evaluation. The inclusion of gender-sensitive indicators and the assurance to mainstream HR and GE across all of the outcomes inevitably depends on in-house capabilities to design the UNDAF results matrix. Furthermore, new Cooperation Frameworks should be linked to national SDG indicator frameworks, which should in turn be linked to regional and global SDG indicator frameworks. Alignment to the SDGs is essential in order to ensure that no one is left behind.

Nevertheless, the UNCTs consulted stated that they often encountered problems in putting together a results framework matrix and lacked sufficient time for this, impacting their ability to include HR and gender-sensitive indicators.

To this end, it is essential that staff are involved in the process from the beginning, and that the design of the indicators and results frameworks is not just an afterthought but is considered one of the most important outputs of the design process. It is also important to ensure that the use of suitable indicators is consulted with the key stakeholders (such as United Nations entities, governments, ministries and civil society). All of these parties should have an active role in developing the indicators, as this will ensure that a twin-track approach is applied to the UNDAF itself and that adequate and consistent monitoring is enabled.

Recommendation Two:

It is important that UNCT staff have the capacity and ability to design HR and GE-responsive results matrices and relevant indicators. Each UNCT should ensure that they either have in-house monitoring and evaluation and expertise in designing gender-responsive results matrices or access to such an expert. There should also be proper accountability and support through the implementation of the UNCT-SWAP Gender Equality Scorecard matrices.

Monitoring and evaluation

While the weak and fragmented design of the results framework is an issue, the monitoring and subsequent collection of data, especially in relation to HR and GE, has been at best weak and on some occasions non-existent. The analysis from the desk review and some of the KIIs revealed that the UNDAF was not always considered 'a living document' and that collaboration and cooperation between agencies on how to communicate results, monitor the work and lobby for resources has at times been fragmented. The onus of responsibility was sometimes misplaced, with some agencies unaware of who should be monitoring which results. There were also incidences in which agencies would provide support to the same ministerial body, thus duplicating efforts. It is

envisaged that the structure and approach towards the new Cooperation Frameworks outlined in the UNDAF guidance will change with the new and reinvigorated Resident Coordinator structure and new generation of Country Teams.

Recommendation Three:

Gender-responsive monitoring and reporting of the results are paramount. The UNCT needs to ensure that monitoring of the results is carried out as stipulated in the UNSDCF guidelines, while ensuring that HR and GE indicators are mainstreamed throughout the results. The accurate and timely collection and monitoring of HR and GE-sensitive data will inevitably contribute to a more HR and GE-responsive evaluation. Efforts should be made to ensure that different groups, such as persons with disabilities and from the LGBTQ+ community, among others, are also disaggregated.

Human rights and gender equality-responsive evaluations

The meta-synthesis used the UN-SWAP EPI scorecard to determine that only four (8 percent) of the 50 evaluations met requirements in terms of their integration of HR and GE. However, a further 27 reports (54 percent) approached requirements, meaning that a total of 31 reports (61 percent) either partially or fully integrated HR and GE into the evaluation reports (finding four).

Quality assurance

The quality of the evaluation report correlates with the extent to which HR and GE perspectives were integrated into the evaluation.

At the time of selecting the sample for the meta-synthesis, only 35 evaluation reports from the time period under evaluation (2015–2019) had been uploaded on the Evaluation Resource Centre portal. Almost all of these had been quality assessed. The meta-synthesis was able to access a further 32 reports which had not been uploaded on the portal or quality assessed. It is unclear how many other UNDAF evaluations from 2015–2019 were carried out and are not publicly accessible.

Finding five concludes that the quality of the evaluation correlates with the extent to which evaluation reports were deemed gender responsive. Reports that tended to be of a higher quality were more likely to integrate HR and GE. The meta-synthesis found that 67 percent of the reports that were rated in quality as moderately satisfactory or above either approached or met UN-SWAP EPI requirements. Quality assurance is paramount to ensure that evaluations meet the required standard and integrate HR and GE.

During evaluation processes, most evaluators received comments and suggestions to improve the report by the UNCT. However, there has been little supportive guidance to align the reports with higher standards and to ensure that these standards conform to the UN-SWAP EPI. Very few reports were critiqued through a gender lens, meaning that many evaluators may be unaware of whether their reports had not sufficiently integrated HR and GE perspectives. It is increasingly becoming normal practice to share checklists as a way of ensuring quality, which provide evaluators with a benchmark against which they can evaluate their performance. This in turn helps evaluators to focus the evaluation and encourages transparency of the process for all.

The meta-synthesis analysed a total of 50 reports, four of which were deemed to be gender responsive. A further 27 reports (54 percent) approached requirements of the UN-SWAP EPI, with many of their weaknesses relating to the evaluators' efforts to integrate HR and GE in the overall scope and design of the evaluation as well as the methodology applied. Had these areas been strengthened, the evaluations would most certainly have scored higher on the UN-SWAP EPI, meaning many of them would most likely have met requirements. To this end, the importance of ToR and its subsequent influence on the design and methodology chosen should not be underestimated.

Recommendation Four:

Inception reports and evaluation reports should be quality assessed through a gender lens using the UN-SWAP EPI criteria, as this will encourage the design and methodology of the evaluation and its subsequent report to be HR and GE responsive. All completed and approved evaluation reports should then be subsequently uploaded to an accessible public repository.

Terms of references

Finding seven concluded that evaluators demonstrated weaknesses in designing GE-responsive evaluations, with 32 percent failing to include GE-responsive questions in the evaluation design. This was sometimes influenced by the fact that the ToR had not included a HR and/or GE scope and lacked any questions focusing on HR and GE.

There are several guidelines⁷³ already in existence which outline tips to ensure that HR and GE perspectives are integrated in the ToR. These include the UNEG guidance on integrating human rights and gender equality in evaluations, published in 2014. Finding seven outlined that 60 percent of the ToR did not follow guidelines and as a result, did not include a HR and GE perspective. Evaluators tended to mirror these ToR and therefore also did not include HR or GE-responsive questions.

Around 57 percent of the ToR examined under this meta-synthesis requested some type of expertise in HR and/or GE, with 85 percent of the evaluation reports either approaching or meeting requirements. Two of the three evaluation reports that missed requirements were deemed unsatisfactory in terms of their quality.

Recommendation Five:

In order to promote a HR and GE-responsive evaluation, UNCTs should use the existing guidelines and templates on drafting ToR to guide them through this process.

ToR should include at a minimum a HR and GE scope and at least one evaluation-related question and should request experience/expertise in HR and GE. UNEG could create gender-responsive templates that can be adapted to different contexts. The final approval of ToR should be overseen by someone who has expertise in HR and GE in order to ensure that these perspectives are integrated in accordance with the context of the Coordination Framework and country.

Quality of evaluators

Under finding five, the meta-synthesis revealed almost 25 percent of the evaluation reports in the sample were of poor quality. Poor-quality evaluations invariably impact on the ability to conduct a HR and GE-responsive evaluation.

Despite not meeting adequate quality assurance standards, the reports had been uploaded and approved. Due to the poor quality of some reports, certain recommendations and findings were not used and the evaluation process was undermined, as a result of the lack of checks and balances required.

Most of the Resident Coordinators consulted stated that it was sometimes difficult to employ good-quality evaluators, as there is a limited number of individuals with experience in evaluating UNDAFs and potentially even less with HR and GE expertise. Employing experienced evaluators is essential in order to ensure a HR and GE-responsive evaluation.

A number of less competent evaluators tend to re-enter the system as there is no blacklisting system in place. As mentioned under recommendation three, not all evaluations are uploaded onto the Evaluation Resource Centre portal and are therefore not quality assessed. RCOs often find themselves with a pool of poor-quality applicants and can do little to improve the quality of the evaluation once an evaluator has been selected.

73

Examples include: UN-Women, *Gender Equality and Big Data* (2018); UNEG, *Good Practices for Integrating Gender Equality and Human Rights in Evaluation* (2017); UN-Women, *How to Manage Gender-responsive Evaluation. Evaluation Handbook* (2015); UNEG, *Integrating Human Rights and Gender Equality in Evaluations* (2014).

While there are a number of weak evaluators, some simply require additional support in order to understand what is expected of them and against which benchmarks the evaluation should be measured. Several guidelines exist on integrating HR and GE into evaluations, though it is important to note that some date back to 2014 and 2015 and do not consider the 2030 Agenda, the inclusion of the SDGs and how evaluations should be conducted.

Although the guidelines are useful, many are too detailed and targeted at United Nations professionals rather than consultants.

Recommendation Six:

RCOs or UNCTs should provide evaluators with support throughout the entire evaluation process to ensure that HR and GE are adequately integrated. The UNEG Evaluation Guidance Compendium states that senior officers should have excellent knowledge of HR and GE and can therefore support the process and other officers overseeing the evaluation. UNEG could provide a report template that includes the mainstreaming of HR and GE throughout the report. Evaluations should be quality assured at all stages through a HR and GE lens, ensuring its compliance with the UN-SWAP EPI. This should be a transparent process that is used as a learning opportunity for evaluators.

Recommendation Seven:

In order to share learning experiences and improve capacities, UNEG could create a roster for evaluators with expertise in HR and GE, which could be shared with country offices in the event that they are unable to find adequate candidates. Furthermore, a public forum for evaluators could be created by the UNEG HR and GE working group, which would enable experienced and non-experienced evaluators to communicate and share their own experiences in conducting a HR and GE evaluation. Examples of gender-responsive reports could also be posted online for evaluators to understand the type of standard that needs to be attained.

While UNEG guidelines are often referred to in evaluations and their respective ToR, they are often not fully used. Despite not always being included in ToR, the UNEG guidance on integrating human rights and gender equality in evaluations is a document that all RCOs should have available to them and they should ensure that staff are familiar with its content.

All the evaluators consulted were aware of the guidelines on integrating HR and GE perspectives, though they rarely used them as a tool. This may be because while useful, the guidelines do not necessarily give specific direction on how to conduct an evaluation of Cooperation Frameworks.

While evaluators should keep to a maximum number of pages for the evaluation reports, UNCTs need to understand that it is not always possible to integrate HR and GE without extending beyond the maximum page count, due to the very short time frames and resources available to carry out the evaluations. Furthermore, time restraints in the field mean that there is often not enough time available for evaluators to cross-check information or carry out a more in-depth assessment of the integration of HR and GE results.

Therefore, support with guidelines and quality assurance is crucial to ensuring that the evaluators are able to produce good-quality reports.

Recommendation Eight:

UNEG should update guidelines by producing a small frequently asked questions document on Cooperation Framework evaluations, aligning this with the 2030 Agenda. UNEG could further produce a PowerPoint tutorial with best practices and sample HR and GE questions for evaluations as well as sample gender-sensitive indicators by using the information already available and summarizing it into one tutorial.

Scope and design

Finding nine concluded that the integration of HR and GE into the scope and overall design of the questions in the evaluation reports was relatively weak, which inevitably led to the evaluations missing requirements.

Only 40 percent of the evaluation reports either satisfactorily or completely integrated a HR and GE scope and questions. In total, 19 reports (38 percent) included a HR or GE scope in the evaluation, with 34 (68 percent) including at least one type of question on HR and GE in the results matrix. However, there were some reports that did not answer these questions in its findings.

All of the reports that had included a HR and/or GE scope also included at least one question on HR and GE. All the reports that included both a HR and GE scope and at least one question either approached or met requirements according to the UN-SWAP EPI.

Only 22 reports (44 percent) included a separate subsection on HR and/or GE, though these sections generally addressed the status of women and did not include information on the results in relation to HR and/or GE. Only four of the reports dedicated an entire section to HR and GE.

In order to ensure that the scope and design of a HR and GE evaluation is adequate, the design should include a number of the aforementioned elements. Unfortunately, only 4 percent (2 reports) included all of these elements, with 50 percent partially integrating them. Given the fact that most evaluators take the lead from ToR, it is essential that ToR integrate a HR and GE perspective. It is particularly important to include different HR and GE-responsive questions under various criteria, especially effectiveness, and where possible encourage a separate section on these issues as well.

Methodology

Finding ten concluded that the second criterion on the UN-SWAP EPI – which examines the extent to which a gender-responsive methodology, methods, tools and data analysis techniques are selected – was also found to be weak, with 30 reports (60 percent) either not integrating any aspect of HR and/GE at all or only partially integrating them.

The meta-synthesis found that that the majority of the reports (94 percent) included a mixed methodology approach, which is customary in HR and GE-responsive evaluations. Nonetheless, only 11 reports (22 percent) included disaggregated data, with much of this cited in the background section of the report. It is assumed that the reluctance of including disaggregated data is partly due to the difficulties encountered by both the Government and the UNCT in collecting relevant data. HR and GE-sensitive indicators were also rare.

Furthermore, only three reports cited in their methodology that it involved consulting individuals outside of the key stakeholders. The UN-SWAP EPI states the importance of ensuring that a wide range of stakeholders is consulted during an evaluation, since this allows for different viewpoints to be obtained as well as an understanding on the extent to which interventions have reached key beneficiaries and groups. Despite this, the normal practice only involved interviewing key stakeholders, including the United Nations agencies, governments and ministries. The evaluations also interviewed civil society groups, though these groups rarely centered around women or vulnerable and marginalized groups.

Evaluators who took part in the survey and KIIs stated that they were often unable to access vulnerable and marginalized groups and that a lack of resources and time to carry out the interviews further hindered their ability to consult with such groups.

Findings, conclusions and recommendations

Finding fourteen revealed that out of 50 reports, a total of 31 reports (62 percent) contained some reference to HR and GE in the background section, 43 reports (86 percent) referred to HR and GE at least once in

the findings and 34 reports (68 percent) included at least one recommendation on HR or GE.

While most evaluation reports referred to HR and GE in their findings (86 percent), they tended not to be mainstreamed throughout the reports, with references included in results primarily found within the focus area of governance.

Impact of the UNDAF on human rights and gender equality

Finding fifteen concluded that the UNDAFs contributed to the promotion of HR and GE, although the real impact is often difficult to gauge. The assessment of the four focus areas revealed that there were references to HR and GE in the programmatic results, 77 percent of which were assessed as gender responsive and therefore examined needs of women, girls and other vulnerable groups.

While the type of results reported most definitely had a positive impact on HR and GE, references tended to focus on one strategic priority, namely governance. Poverty, economic empowerment and the environment and climate change all had less references to HR and GE, which were also less gender responsive. These findings tie in with the recommendations expressed in the Secretary-General's report on mainstreaming a gender perspective into all policies and programmes in the United Nations system, which emphasized the need to increase "efforts to integrate a gender perspective in non-traditional technical areas of work such as infrastructure and energy to ensure that related policies and programmes address the needs and priorities of women and girls."⁷⁴

Key drivers of change for human rights and gender equality

Finding sixteen concluded that the evaluations' results contributed equally towards improving gender-sensitive policies and laws and enabling access to resources and opportunities. The different areas tended to use different spheres or levels of change to obtain their results.

Governance and climate change would use the change in policies and laws to drive most of their changes, while economic empowerment and poverty focused the majority of their efforts on access to resources and opportunities. All four areas paid little attention to internal culture and exclusionary practices as ways in which to harness transformative change.

While many of the results were found to be gender responsive, transformative changes were only seen under nine results. However, most of these results had the potential to lead to transformative change, though more time and a more strategic approach may be required to ensure adequate gender-responsive monitoring and reporting of the results.

Addressing the leave no one behind principle

The meta-synthesis found that there were more references to gender-responsive results than HR results. Some offices reported difficulties in promoting HR due to the social, political and economic context of the country, which impacted on the extent to which HR results were mainstreamed across the outcomes.

Finding seventeen concluded that there was little mention of vulnerable groups and that women still feature very predominantly. Youth were targeted especially under economic empowerment and governance but did not feature much in most results.

Several other groups, such as persons with disabilities and refugees, were also mentioned, though to a far lesser extent. In order for any future Cooperation Frameworks to address discrimination, exclusion, inequalities and vulnerabilities, which can impact on any one of the SDGs, it is essential that results reach other groups. This requires using disaggregated data and qualitative analysis to identify who is being excluded or discriminated against along with how and why this is happening, as well as who is experiencing multiple and intersecting forms of discrimination and inequalities. To this end, the weakness in data-collection and reporting needs to be addressed.

The efficiency of inter-agency cooperation to contribute to human rights and gender equality results

Finding eighteen found it difficult to ascertain the extent to which inter-agency cooperation contributed to greater efficiency on HR and GE, as very few reports referred to HR and GE when assessing inter-agency cooperation.

The evaluation centered around strengthening monitoring and evaluation frameworks, establishing more fluid communication between the agencies and improving the coordination of working groups and joint resource mobilization.

The implementation of the UNCT-SWAP Gender Equality Scorecard can revitalize the Gender Theme Groups and substantially contribute to enhancing gender mainstreaming at the institutional and programmatic levels in the UNCTs.

It is envisaged that inter-agency cooperation will be strengthened by Cooperation Framework guidelines and that agencies will be better able to work together. Furthermore, approximately 30 percent of the evaluation reports referred to the involvement of other actors in implementation and monitoring processes, which is a particularly important role for civil society.

Recommendation Nine:

UNCTs should develop a joint gender strategy to ensure that there is a consistent and aligned approach for monitoring results, holding HR and GE working groups and ensuring adequate resource mobilization are met, utilizing the implementation of the UNCT-SWAP Gender Equality Scorecard.⁷⁵

75

²⁰. Gender equality and women's empowerment are integral to realizing the 2030 Agenda and all of the SDGs. To integrate a focus on these issues throughout the Cooperation Framework, UN development entities should put gender equality at the heart of programming, driving the active and meaningful participation of both women and men, and consistently empowering women and girls, in line with the minimum requirements agreed upon by the United Nations Sustainable Development Group (UNSDG) in the UNCT System-wide Action Plan (SW/AP) Gender Equality Scorecard." UNSDG, *United Nations Sustainable Development Cooperation Framework. Internal Guidance* (2019). Available at https://undg.org/wp-content/uploads/2019/06/UN-Cooperation-Framework-Internal-Guidance-25_June-2019.pdf.

VI. Annexes

ANNEX I. Overall assessment of the fifty evaluations under the UN-SWAP EPI

APPROACHES	COUNTRY	INCOME STATUS	LANGUAGE	OVERALL RATING	OVERALL SCORE	CRITERION 1	CRITERION 2	CRITERION 3
Requirements	Ecuador	UMIC	Spanish	Approaches Requirements	5	2	2	1
Asia & the Pacific	Viet Nam	LMIC	English	Approaches Requirements	5	1	2	2
Africa	Guinea	LMIC	French	Approaches Requirements	5	2	2	1
Africa	Côte D'Ivoire	LMIC	French	Misses Requirements	3	1	1	1
Asia & the Pacific	Bhutan	UMIC	English	Approaches Requirements	6	2	2	2
CIS	Moldova	LMIC	English	Meets Requirements	8	2	3	3
Asia & the Pacific	Papua New Guinea	LMIC	English	Approaches Requirements	4	2	1	1
Africa	Mozambique	LIC	English	Approaches Requirements	6	3	2	1
Africa	Botswana	UMIC	English	Approaches Requirements	4	1	1	2
CIS	Albania	UMIC	English	Approaches Requirements	5	2	1	2
Africa	Senegal	LIC	English	Misses Requirements	3	0	1	2
Africa	Togo	LIC	French	Meets Requirements	7	2	2	3
Latin America & the Caribbean	Belize	UMIC	English	Misses Requirements	3	1	1	1
Africa	Cape Verde	LMIC	English	Approaches Requirements	5	2	2	1
Africa	Gambia	LIC	English	Approaches Requirements	5	2	1	2
Asia & the Pacific	Bangladesh	LMIC	English	Approaches Requirements	6	2	2	2
Latin America & the Caribbean	Guyana	UMIC	English	Misses Requirements	3	1	1	1

APPROACHES	COUNTRY	INCOME STATUS	LANGUAGE	OVERALL RATING	OVERALL SCORE	CRITERION 1	CRITERION 2	CRITERION 3
Latin America & the Caribbean	Honduras	LMIC	Spanish	Approaches Requirements	6	3	2	1
Africa	Sudan	LMIC	English	Misses Requirements	3	1	1	1
Latin America & the Caribbean	El Salvador	LMIC	Spanish	Misses Requirements	3	1	1	1
Africa	Eswatini	LMIC	English	Misses Requirements	3	1	1	1
Latin America & the Caribbean	Guatemala	UMIC	Spanish	Approaches Requirements	6	2	2	2
Africa	Cameroon	LMIC	French	Approaches Requirements	4	1	2	1
Latin America & the Caribbean	Chile	HIC	Spanish	Misses Requirements	1	0	0	1
Africa	Nigeria	LMIC	English	Misses Requirements	3	1	1	1
CIS	Turkey	UMIC	English	Misses Requirements	3	1	1	1
Latin America & the Caribbean	Cuba	UMIC	Spanish	Misses Requirements	2	1	1	0
Africa	Central African Republic	LIC	French	Misses Requirements	2	0	1	1
Africa	Congo	LIC	French	Approaches Requirements	4	1	1	2
Latin America & the Caribbean	Costa Rica	UMIC	Spanish	Misses Requirements	3	1	1	1
Arab States	Egypt	LMIC	English	Approaches Requirements	4	1	1	2
Africa	Chad	LIC	French	Misses Requirements	2	0	1	1
Asia & the Pacific	Samoa	UMIC	English	Approaches Requirements	6	2	2	2
Africa	Ghana	LMIC	English	Approaches Requirements	4	1	1	2
Asia & the Pacific	Afghanistan	LIC	English	Misses Requirements	2	1	0	1

APPROACHES	COUNTRY	INCOME STATUS	LANGUAGE	OVERALL RATING	OVERALL SCORE	CRITERION 1	CRITERION 2	CRITERION 3
Arab States	Iran	UMIC	English	Misses Requirements	2	1	0	1
Arab States	Jordan	UMIC	English	Misses Requirements	3	1	1	1
Africa	Kenya	LMIC	English	Approaches Requirements	5	1	2	2
CIS	Kyrgyzstan	LMIC	English	Meets Requirements	8	2	3	3
Africa	Liberia	LIC	English	Meets Requirements	7	2	2	3
CIS	Montenegro	UMIC	English	Approaches Requirements	5	1	2	2
Arab States	Palestine	LMIC	English	Misses Requirements	3	1	1	1
Africa	Philippines	LMIC	English	Approaches Requirements	5	1	2	2
Africa	Rwanda	LIC	English	Approaches Requirements	5	1	2	2
Africa	Somalia	LIC	English	Approaches Requirements	5	2	1	2
Africa	South Sudan	LIC	English	Approaches Requirements	4	2	1	1
Africa	Sri Lanka	LMIC	English	Misses Requirements	1	0	0	1
CIS	Ukraine	LMIC	English	Approaches Requirements	4	1	1	2

Notes: LIC = low-income country; LMIC = lower-middle-income country ; UMIC = upper-middle-income country; HIC = high-income country.

ANNEX II. Bibliography

Abreu Lopes, Claudia, and Bailur, Savita. *Gender Equality and Big Data: Making Gender Data Visible. UN-Women Innovation Facility*, 2018.

Acón Monge, Karol, and Garnier Rímolo, Leonardo. *Evaluación del Marco de Asistencia de Naciones Unidas para el Desarrollo 2013–2017 [Evaluation of the United Nations Development Assistance Framework 2013–2017]*. Office of the United Nations System Resident Coordinator in Costa Rica, 2016.

Antelo Seborga, Natalie. *Informe de evaluación. Consultoría para evaluar el Marco de Asistencia para el Desarrollo del Sistema de Naciones Unidas en Chile [Evaluation Report. Consultation to Evaluate the United Nations Development Assistance Framework in Chile]*. United Nations Chile, 2018.

Arbulu, Angelica, Lindström, Johanna, Quang Nam, Pham, and Thi Mai Huong, Le. *Independent Review of the One Plan (2012–2016) between the Government of the Socialist Republic of Viet Nam and the United Nations in Viet Nam. Final Report*. Swedish Institute for Public Administration (SIPU), 2016.

Barnes, Joseph. *Good Practices for Integrating Gender Equality and Human Rights in Evaluation*. United Nations Evaluation Group, 2017.

Bester, Angela. *Terminal Report: An Assessment of the United Nations Development Assistance Framework (UNDAF) for Swaziland, 2011–2015. Final Report*. United Nations Eswatini, 2015.

BetterEvaluation. Most Significant Change. n.d. https://www.betterevaluation.org/en/plan/approach/most_significant_change.
Cape Verde, Ministry of Foreign Affairs, and United Nations Cape Verde. *Plano operacional do UNDAF 2012–2016 [UNDAF Action Plan 2012–2016]*. Ministry of Foreign Affairs and United Nations Cape Verde, n.d.

Carrondo, Sofia. *Evaluation UN(D)AF 2013–2017. Hashemite Kingdom of Jordan*. United Nations Jordan, 2016.

Chiwara, Richard, Aldar, Dolgion, and Sergelen, Battulga. *Evaluation of the United Nations Development Assistance Framework for Mongolia UNDAF (2012–2016)*. Government of Mongolia and United Nations Mongolia, 2015.

Faye, Cheikh, and Nzukam, Guillaume. *Evaluation Final UNDAF Cycle 2013–2017. Rapport Final Principal [Final Evaluation UNDAF Cycle 2013–2017. Main Final Report]*. United Nations Cameroon, 2018.

Food and Agricultural Organization of the United Nations (FAO). *Pacific Multi-Country CPF Document 2013–2017*. FAO, 2012.

Gallicchio, Enrique and Hernández, Alina. *Informe de Evaluación – Evaluación final – Marco de Asistencia de las Naciones Unidas para el Desarrollo en Cuba (2014–2019) [Evaluation Report – Final Evaluation – United Nations Development Assistance Framework in Cuba (2014–2019)]*. United Nations Cuba, n.d.

Geoffrey, Batali. *Evaluation of South Sudan Interim Cooperation Framework (ICF) (2016–2018). Evaluation Report*. United Nations South Sudan, 2018.

Góchez Rebollo, Mercedes. *Evaluación del Marco de Asistencia de las Naciones Unidas para el Desarrollo – UNDAF 2012–2015. El Salvador [Evaluation of the United Nations Development Assistance Framework – UNDAF 2012–2015. El Salvador]*. United Nations El Salvador, 2015.

Government of Cameroon and United Nations Cameroon. *UNDAF 2013–2017. Pour une croissance inclusive et la protection des personnes vulnérables [UNDAF 2013–2017. For Inclusive Growth and Protection of Vulnerable People]*. United Nations Cameroon, n.d.

Government of Congo and United Nations Congo. *Plan Cadre des Nations Unies pour L'Aide au Développement 2014–2018 [United Nations Development Assistance Framework 2014–2018]*. Government of Congo and United Nations Congo, 2013.

Government of the Republic of Honduras and United Nations Honduras. *Plan de Acción del Marco de Asistencia de las Naciones Unidas para el Desarrollo [Action Plan of the United Nations Development Assistance Framework 2012–2016]*. United Nations Honduras, 2011.

Government of the Republic of Turkey and United Nations Turkey. *United Nations Development Cooperation Strategy Turkey 2011–2015*. Ankara: Government of the Republic of Turkey and United Nations Turkey, 2010.

Government of Sudan. *United Nations Development Assistance Framework for the Republic of Sudan*. Khartoum, 2012.

Government of Togo and United Nations Togo. *Plan Cadre des Nations Unies pour l'Aide au Développement (UNDAF) 2014–2018 [United Nations Development Assistance Framework (UNDAF) 2014–2018]*. United Nations Togo, 2013.

Government of Ukraine and United Nations Ukraine. *Government of Ukraine-United Nations Partnership Framework 2012–2016*. Kiev: Government of Ukraine and United Nations Ukraine, 2011.

Hagona, Elballa. *Evaluation Report of Sudan UNDAF (2013–2016)*. United Nations Sudan, 2015.

Husain, Saiyed M. *United Nations Development Assistance Framework Afghanistan (2010–2014). Final Report*. United Nations Afghanistan, 2015.

INTRAC. *Outcome harvesting*. 2017. <https://www.intrac.org/wpcms/wp-content/uploads/2017/01/Outcome-harvesting.pdf>.

Ittig, Annette. *The 2012–2016 Gambia United Nations Development Assistance Framework Final Evaluation Report*. United Nations Gambia, 2016.

Ittig, Annette, and Karanja, Michael. *United Nations Development Assistance Framework 2014–2018. Final Evaluation Report*.

United Nations Kenya, 2017.

Jambrina Rodríguez, Fernando. *Final Evaluation Report. United Nations Development Assistance Framework in Honduras 2012–2016*. Office of the United Nations System Resident Coordinator in Honduras, 2015.

Kabore, Roger Bila, and Guilavogui, Massa. *Evaluation Final du Plan Cadre des Nations Unies pour l'Aide au Développement de la Guinée (UNDAF 2013–2017) [Final Evaluation of the UN Framework Plan for Development Assistance to Guinea (UNDAF 2013–2017)]*. United Nations Guinea, 2017.

Kam, Oleh, Assogba, Kouassi, Tekou, Koffi Bléwussi and Aguey, Kpati (2017). *Évaluation finale du Plan Cadre des Nations Unies pour l'aide au Développement (UNDAF) 2014–2018. Rapport provisoire [Final Evaluation of the United Nations Development Assistance Framework (UNDAF) 2014–2018. Draft Report]*. United Nations Togo.

Kouidhi, Moncef, Ahoure, Alban and Dogore, Raymond. *UNDAF 2013–2015. Evaluation final. Acquis, leçons apprises et orientations stratégiques pour 2017–2020. Rapport final [UNDAF 2013–2015. Final Assessment. Knowledge, Lessons Learned and Strategic Approaches for 2017–2020. Final Report]*. United Nations Côte d'Ivoire, 2015.

Kouidhi, Moncef, and Boungou, Jean-Christophe. *UNDAF 2014–2018. Evaluation. Acquis, leçons apprises et orientations stratégiques et programmatiques [UNDAF 2014–2018. Evaluation. Knowledge, Lessons Learned and Strategic and Programmatic Approaches]*. United Nations Congo, n.d.

KPMG. *Evaluation of UNDAF 2012–2016. Final Report*. United Nations Mozambique, 2015.

Long, Soraya, Arce, Natalie, Kristensen, Christel, Centeno, Henry, and Mayorga, Dulce María. *Evaluación final del Marco de Asistencia de Naciones Unidas para el Desarrollo en Guatemala: 2015–2019 (UNDAF) [Final Evaluation of the United Nations Development Assistance Framework in Guatemala: 2015–2019 (UNDAF)]*. United Nations Guatemala, n.d.

- Lund, Ann. *Independent Evaluation. UN Development Assistance Framework (2013–2017)*. United Nations Country Team in the Pacific, 2016.
- Lund, Ann, and Lovai, Betty. *Independent Evaluation. UN Development Assistance Framework (2012–2017)*. United Nations in Papua New Guinea, 2016.
- Maccelli, Donata. *Belize UNDAF 2013–2016 Evaluation. Final Report*. United Nations Belize, 2017.
- MacKenzie, Alexander, and Ymeri, Sabina. *Evaluation of the Government of Albania and United Nations Programme of Cooperation (PoC) 2012–2016: Final Report*. United Nations Albania, 2015.
- Mbodji, Serigne, and Nignon, Pierre. *Évaluation finale UNDAF 2014–2018. Rapport final [Final Evaluation UNDAF 2014–2018. Final Report]*. United Nations Niger, 2018.
- Mikayoulou, Daniel and Manga Badji, Anne Nathalie. *Plan Cadre des Nations Unies pour l'Aide au Développement du Senegal 2012–2018. Rapport d'évaluation finale [United Nations Development Assistance Framework 2012–2018 for Senegal. Final Evaluation Report]*. United Nations Senegal, n.d.
- Mpswenumugabo, Théodore, Keïta, Karounga, Moyangar, Toralta Tell and Diallo, Aly. *Rapport d'évaluation finale du cadre stratégique interimaire CSI 2012–2015 [Final Evaluation Report of the Interim Strategic Framework 2012–2015]*. United Nations Chad, 2016.
- Naumann, Craig. *Final Report. Evaluation of Cabo Verde's United Nations Development Assistance Framework (UNDAF) 2012–2017*. United Nations Cape Verde, 2016.
- Naumann, Craig. *Somalia ISF Review – Final Report*. United Nations Somalia, 2016.
- Ngatoua, Paulin Calixte. *Évaluation Finale du Cadre Stratégique Intérimaire du Système des Nations-Unies en RCA CSI 2016–2017. Rapport Final. [Final Evaluation of the Interim United Nations Strategic Framework in CAR 2016–2017. Final Report]*. United Nations Central Africa Republic, 2018.
- Norwegian Agency for Development Cooperation (Norad). *Gender and Development – A Review of Evaluation Reports 1997–2004. Evaluation Report 3/2005*. Oslo, 2005.
- Norwegian Agency for Development Cooperation (Norad). *Lessons from Evaluations of Women and Gender Equality in Development Cooperation. Synthesis Report 2006/1*. Oslo, 2006.
- Novillo, Nathalia. *Informe de la evaluación final del Marco de Cooperación de las Naciones Unidas en Ecuador 2015–2018 [Report of the final evaluation of the United Nations Cooperation Framework in Ecuador 2015–2018]*. United Nations Ecuador, 2017.
- Nuñez Sabaris, Alberto, Cristina Udelsmann Rodrigues, and Carlos Manuel López. *Avaliação do UNPAF 2015–2019 para Angola. Relatório Final [Angola UNDAF 2015–2019 Evaluation. Final Report]*. United Nations Angola, 2018.
- Nuwakora, Cliff Bernard. *Independent Evaluation of the Govt of Botswana – UN Programme Operational plan 2010–2016. Final Report*. United Nations Botswana, 2015.
- Nuwakora, Cliff Bernard and Nicholas Guribie. *UNDAF 2012–2017 Evaluation Report. Final Draft Report*. United Nations Ghana, 2016.
- Olver, Richard. *United Nations Partnership Framework (UNPF) 2012–2017. Forward Looking Evaluation and Comparative Advantage Analysis*. United Nations in Ukraine, 2016.
- Orlandini, Barbara. *Independent Evaluation. UNDAF 2012–16 Guyana. Final Report*. United Nations in Guyana, 2015.
- Piga, Graziella, Novovic, Tomislav and Mogileskii, Roman. *Final Evaluation of the United Nations Development Assistance Framework (UNDAF) for Kyrgyz Republic 2012–2017*. United Nations Kyrgyz Republic, 2016.

Pittman, Alexandra. *Fast-Forwarding Gender Equality and Women's Empowerment? Reflections on Measuring Change for UNDP's Thematic Evaluation on Gender Mainstreaming and Gender Equality 2008–2013*. Independent Evaluation Office Occasional Paper. New York: UNDP, Independent Evaluation Office, 2014.

Privat, Christian. *Evaluation of the Integrated UN Programme. United Nations Development Assistance Framework in Montenegro (2012–2016)*. United Nations in Montenegro, 2015.

Seamus, C., A. Onwumele, A. and Smah, S. *End of Cycle Evaluation of UN Development Assistance Framework (UNDAF) III, 2014–2017. Final Report*. United Nations Nigeria, 2018.

Sibanda, Backson, Matshalaga, Neddy, Bloh, O., Hove, T., Maroveke, Macreynold, Turugari, Leonard and Mtawarira, D. *Liberia UNDAF 2013–2017. Evaluation Report. United Nations Development Assistance Framework (UNDAF) Liberia 2013–2017*. Primson Management Services, 2018.

UDA Consulting. *Evaluation of the UNDCS for Turkey (2011–2015) – Final Report*, 2016.

United Fund. *Companion Guide Evaluation Phase. SR Edits 310719*. United Fund, n.d.

UN Women. *Guidance Note. Gender Mainstreaming in Development Programming*. New York: UN System Coordination Division of UN Women, 2014.

UN Women. *How to Manage Gender-Responsive Evaluation. Evaluation Handbook*. New York: Independent Evaluation Office, 2015.

UN Women. *TRANSFORM, vol. 1* (June 2015).

United Nations. Meta-evaluation and synthesis of United Nations Development Assistance Framework evaluations, with a particular focus on poverty eradication (JIU/REP/2016/6). Supplementary paper to JIU/REP/2016/6. 2006.

United Nations, Economic and Social Council. Mainstreaming a gender perspective into all policies and programme in the United Nations system. Report of the Secretary-General. 6 April 2017. E/2017/57.

United Nations, General Assembly. 70/1. Transforming our world: the 2030 Agenda for Sustainable Development. 21 October 2015. A/RES/70/1.

United Nations, Joint Inspection Unit. Meta-evaluation and synthesis of United Nations Development Assistance Framework evaluations, with a particular focus on poverty eradication. 2006. JIU/REP/2016/6.

United Nations Albania. *Government of Albania and United Nations. Programme of Cooperation 2012–2016*. United Nations Albania, n.d.

United Nations Angola. *Partnership Framework between the Government of Angola and the United Nations System (UNPAF) 2015–2019*. United Nations Angola, n.d.

United Nations Belize. *Memorandum of Understanding. United Nations Development Assistance Framework 2013–2016*. United Nations Belize, n.d.

United Nations Bhutan. *United Nations Development Assistance Framework One Programme (2014–2018) Evaluation. Gross National Happiness Commission and the United Nations Country Team Bhutan*, 2018.

United Nations Botswana. *Government of Botswana and United Nations Sustainable Development Framework (UNSDF) 2017–2021*. United Nations Botswana, n.d.

United Nations Chile. *Marco de Asistencia para el Desarrollo del Sistema de Naciones Unidas en Chile 2015–2018 [United Nations Development Assistance Framework in Chile 2015–2018]*. United Nations Chile, 2014.

United Nations Development Group (UNDG). *Gender Equality Marker. Guidance Note*. New York: United Nations Development Operations Coordination Office, 2013.

United Nations Development Group (UNDG). *Theory of Change. UNDAF Companion Guidance*. UNDG, 2017.

United Nations Development Group (UNDG). *UN Vision 2030. UNDAF Companion Guidance*. UNDG, 2017.

United Nations Development Group (UNDG). *United Nations Development Assistance Framework Guidance*. United Nations Development Operations Coordination Office, 2017.

United Nations Development Group (UNDG). *Resource Book for Mainstreaming Gender in UN Common Programming at the Country Level*. UNDG, 2018.

United Nations Development Operations Coordination Office. *Gender and the UNDS Reform. Gender in the New Cooperation Framework*. United Nations Development Coordination Office, n.d.

United Nations Development Operations Coordination Office. *2016 United Nations Development Assistance Framework. Desk Review*. United Nations Development Operations Coordination Office, 2016.

United Nations Development Operations Coordination Office. *Building Blocks towards 2030. UNDG Standard Operating Procedures for 'Delivering as One' 2015 Progress Report*. United Nations Development Operations Coordination Office, 2016.

United Nations Development Operations Coordination Office. *Summary brief on UN Development Assistance Frameworks (UNDAFs) – Status, trends and next generation*. United Nations Development Operations Coordination Office, 2016.

United Nations Development Programme (UNDP). *Gender Equality. Evaluation of Gender Mainstreaming in UNDP*. New York: UNDP, 2006.

United Nations Development Programme (UNDP). *Evaluation of UNDP Contribution to Gender Equality and Women's Empowerment*. New York: UNDP, Independent Evaluation Office, 2015.

United Nations Development Programme (UNDP). *Summary: Evaluation of UNDP Contribution to Gender Equality and Women's Empowerment*. New York: UNDP, Independent Evaluation Office, 2015.

United Nations Development Programme (UNDP). *What does it mean to leave no one behind? A UNDP discussion paper and framework for implementation*. New York: UNDP, 2018.

United Nations Development Programme (UNDP). *UNDP Evaluation Guidelines*. New York: Independent Evaluation Office of UNDP, 2019.

United Nations Ecuador. *Marco de Cooperación de las Naciones Unidas en Ecuador 2015–2018 [United Nations Cooperation Framework in Ecuador 2015–2018]*. United Nations Ecuador, 2014.

United Nations Egypt. *United Nations Development Assistance Framework for Egypt (2013–2017). Achieving MDGs+ with Inclusive Growth, Freedom, Social Justice and Dignity*. United Nations Egypt, n.d.

United Nations El Salvador. *Marco de Asistencia de Naciones Unidas para el Desarrollo 2012–2015 [United Nations Development Assistance Framework 2012–2015]*. United Nations El Salvador, 2011.

United Nations Europe and Central Asia Issue-Based Coalition on Gender Equality for Europe and Central Asia. *Desk Review of 18 UNDAFs in Europe and Central Asia. Application of Twin-Track Approach to Gender in UNDAFs and CCAs to Promote Gender Equality and Empowerment of Women*. United Nations Development Group (UNDG), 2018.

United Nations Evaluation Group. *Integrating Human Rights and Gender Equality in Evaluations*. Sales No. E.14.III.B.18.

United Nations Evaluation Group. *Frequently Asked Questions for UNDAF Evaluations*. United Nations Evaluation Group, 2011.

United Nations Evaluation Group. *Guidance on Evaluating Institutional Gender Mainstreaming*. United Nations Evaluation Group, 2018.

United Nations Evaluation Group. *UN-SWAP Evaluation Performance Indicator Technical Note*. United Nations Evaluation Group, 2018.

United Nations Gambia. *United Nations Development Assistance Framework 2012–2016*. United Nations Gambia, 2011.

United Nations Ghana. *Action Plan for the United Nations Development Assistance Framework (UNDAF) 2012–2016*. United Nations Ghana, 2011.

United Nations Guatemala. *Marco de Asistencia de la Naciones Unidas para el Desarrollo –UNDAF– 2010–2014 [United Nations Development Assistance Framework –UNDAF– 2015–2019]*. United Nations Guatemala, 2014.

United Nations Guyana. *United Nations Development Assistance Framework 2013–2016*. United Nations Guyana, 2011.

United Nations Islamic Republic of Iran. *United Nations Development Assistance Framework. The Islamic Republic of Iran 2012–2016*. United Nations Islamic Republic of Iran, 2011.

United Nations Jordan. *United Nations Development Assistance Framework 2013–2017*. United Nations Jordan, n.d.

United Nations Kenya. *United Nations Development Assistance Framework (UNDAF) 2018–2022*. United Nations Kenya, n.d.

United Nations Liberia. *One Programme. United Nations Development Assistance Framework 2013–2017*. United Nations Liberia, 2013.

United Nations Moldova. *Final Evaluation of UN-Republic of Moldova Partnership Framework 2013–2017. Final Evaluation Report (vol.1)*. United Nations Republic of Moldova, 2016.

United Nations Moldova. *National Consultant for the Final Evaluation of the UN-Republic of Moldova Partnership Framework 2013–2017 (UNPF)*. United Nations Moldova, 2016.

United Nations Mozambique. *United Nations Assistance Development Framework for Mozambique 2012–2015 Action Plan*. United Nations in Mozambique, 2011.

United Nations Niger. *Plan Cadre des Nations Unies pour l'Assistance au Développement (UNDAF). Niger 2014–2018 [United Nations Development Assistance Framework (UNDAF). Niger 2014–2018]*. United Nations in Niger, n.d.

United Nations Office at Vienna and United Nations Office on Drugs and Crime. *Strategy for Gender Equality and the Empowerment of Women (2018–2021)*. Vienna: United Nations Office at Vienna, 2018.

United Nations Papua New Guinea. *United Nations Development Assistance Framework (UNDAF) 2012–2015*. Papua New Guinea. United Nations Papua New Guinea, 2012.

United Nations Philippines. *Supporting Inclusive, Sustainable and Resilient Development. The United Nations Development Assistance Framework for the Philippines 2012–2018*. Makati City: UN Coordination Office, 2011.

United Nations Philippines. *Philippines United Nations Development Assistance Framework 2012–2018. Evaluation Report*. United Nations in Philippines, 2017.

United Nations Rwanda. *Rwanda United Nations Development Assistance Plan 2013–2018*. United Nations in Rwanda, 2013.

United Nations Somalia. *United Nations Somalia Assistance Strategy 2011–2015*. United Nations in Somalia, n.d.

United Nations Sri Lanka. *United Nations Development Assistance Framework (UNDAF) (2013–2017)*. Colombo: United Nations in Sri Lanka, 2012.

United Nations Sustainable Development Group (UN SDG). *UNCT-SWAP Gender Equality Scorecard. Performance Indicators for Gender Equality and the Empowerment of Women for United Nations Country Teams. 2018 Framework & Technical Guidance*. New York: United Nations Development Operations Coordination Office, 2018.

United Nations Sustainable Development Group (UN SDG). *United Nations Sustainable Development Cooperation Framework. Internal Guidance*. UN SDG, 2019.

United Nations System Chief Executives Board for Coordination. *Leaving No One Behind: Equality and Non-Discrimination at the Heart of Sustainable Development. A Shared United Nations System Framework for Action*. New York: United Nations, 2017.

Universalia Management Group. *Final Evaluation of UNDAF Rwanda 2013–2018, 2017*.

Verité Research. *Final Evaluation of the United Nations Development Assistance Framework (UNDAF) (2013–2017). Final Report*. United Nations Sri Lanka, 2017.

