Challenges of integrating gender equality in evaluation
# CONTENTS

1.0 INTRODUCTION 7

2.0 BACKGROUND 8
2.1 Gender equality
2.2 What is gender equality in evaluation?

3.0 UNITED NATIONS (UN) WORK ON GENDER EQUALITY IN EVALUATION 10
3.1 UN System Wide Action Plan (UN-SWAP) for gender equality and the empowerment of women 10
3.2 UN-SWAP Evaluation Performance Indicators 11
3.3 Integrating Human Rights and gender Equality in Evaluation – Towards UNEG Guidance 12

4.0 CHALLENGES OF INTEGRATING GENDER EQUALITY IN EVALUATION 14
4.1 Conceptual difficulties 14
4.2 Methodological difficulties 15
4.3 Framework difficulties 16

5.0 HOW UN WOMEN IS OVERCOMING THE CHALLENGES 18
5.1 Strengthen gender sensitive indicators 19
5.2 Planning evaluations in programme cycle 20
5.3 Methodological improvement 20
5.4 Evaluation capacity development 21

6.0 CONCLUSIONS 22

REFERENCES 23
This paper examines the current developments of gender dimensions in evaluation in the United Nations (UN) system. The United Nations Evaluation Group (UNEG), which brings together all the evaluation offices of the UN, has played an important role in setting the evaluation Norms and Standards to highlight the need for people centered evaluation. These initiatives include the development of UN System Wide Action Plan (UN-SWAP) for gender equality and the empowerment of women, UN-SWAP Evaluation Performance Indicators, and the Guidance of Integrating Human Rights and Gender Equality and Evaluation –Towards UNEG. These documents are to provide step by step guidance on how to integrate human rights and gender equality throughout an evaluation process, and are indeed influencing on the way the UN is conducting a gender responsive evaluation. Yet, applying the Guidance is not always easy and the paper will examine those difficulties, in terms of concept, methodology and framework, using our real evaluation management experiences in UN Women Asia and the Pacific region. In addition, UN Women Asia and the Pacific would like to introduce how we are trying to overcoming the challenges based on our practical experiences in the region.

KEYWORDS
GENDER, EVALUATION, UN, SWAP, UNEG GUIDANCE, CONCEPTS, METHODOLOGIES, FRAMEWORKS
INTRODUCTION

“GENDER” HAS BECOME an integral part of development thinking and practices in the last three decades. Women-specific development initiatives, special gender sensitization programmes, and “gender” components in ongoing development projects have been launched by government and non-government organizations, and are supported by international development organizations and donor agencies. These development initiatives and programmes have been accompanied by the thrust towards measuring, monitoring and evaluating the success of the programmes in changing women’s realities and transforming gender relations, in short, what is known as “gender evaluation”. In this paper, first, UN Women Asia and the Pacific will explain what gender equality in evaluation and why it matters.

Second, we will examine how the United Nations (UN), particularly the United Nations Evaluation Development Group (UNEG), which is a professional network that brings together the units responsible for evaluation in the UN system, have been working on gender equality dimensions in evaluation. These initiatives include the development of UN System Wide Action plan (UN-SWAP) for gender equality and the empowerment of women, UN-SWAP Evaluation Performance Indicators, and Integrating Human Rights and Gender Equality in Evaluation – Towards UNEG Guidance.

Third, we will analyze challenges of applying the concept and the Guidance of gender equality to evaluation. According to the UNEG Guidance on Integrating Human Rights and Gender Equality in Evaluation, the principles of Human Rights and Gender Equality (HR & GE) in evaluation are: Inclusion, Participation, Fair Power Relations and Mixed Methods. I will examine the challenges of conducting gender sensitive evaluation in terms of conceptual, methodological and framework application.

Finally, we will explain how UN Women is trying to overcome the challenges in order to make our evaluations more gender responsive. There are many innovations in the organization, and in this section, we would like to introduce four of our practical solutions – such as strengthening gender sensitive indicators, planning evaluations’ types and timing in the 3-5 years programme cycle, improving methodologies and building evaluation capacities for staff and partners in innovative ways.

Throughout the paper, the current situation of gender equality in evaluation in the UN, its challenges, and the initiatives to overcome the challenges will be discussed.
Challenges of Integrating Gender Equality in Evaluation

UN Women/Betsy Davis
BACKGROUND

2.1 Gender equality

GENDER EQUALITY REFERS to the equal rights, responsibilities and opportunities of women and men. Gender equality does not imply that women and men are the same, but they have equal value and should be accorded equal treatment.

The United Nations (UN) regards gender equality as a human right. It points out that empowering women is also an indispensable tool for advancing development and reducing poverty.

The 1979 Convention on the Elimination of All Forms of Discrimination against Women (CEDAW) led the UN and its members to stress the importance of operationalizing gender equality. The Beijing Platform of Action and the Millennium Declaration also commit the UN to promoting gender equality in its development efforts. 1

Despite this, gender equality is still hard to achieve.

For instance, equal payment for equal work is one of the areas where gender equality is rarely seen. All too often women are paid less than men for doing the same work. This is one of the reasons that the majority of the world's poor are women: around 70% of the people who live in extreme poverty, on less than US$ 1 a day, are girls and women.

Furthermore, discrimination against women and girls, such as gender-based violence (GBV), economic discrimination, reproductive health inequalities and harmful traditional practices, remains the most pervasive and persistent form of inequality.

2.2 What is gender equality in evaluation?

“GENDER” HAS BECOME an integral part of development thinking and practices in the last three decades. Women-specific development initiatives, special gender sensitization programmes, and “gender” components in ongoing development projects have been launched by government and non-government organizations, and are supported by international development organizations and donor agencies. These development initiatives and programmes have been accompanied by the thrust towards measuring, monitoring and evaluating the success of the programmes in changing women’s realities and transforming gender relations, in short, what is known as “gender evaluation”. 2

Although much of the literature on gender evaluation continues to emerge from scholars and development practitioners, the terminologies are still diversified: “gender equality in evaluation”, “gender evaluation”, “gender sensitive evaluation”, “gender responsive evaluation”, “feminist evaluation”, “engendered evaluation” etc. 3

In the absence of the standard terminology and conceptual framework, we would like to use the language “gender equality in evaluation” in this paper, and define it as follows:

“Gender equality in evaluation” involves learning how change happens in gender relations, analyzing which strategies worked and which did not, and how these could be refined for greater impact, practicing accountability among the various stakeholders to empowering constituencies by involving them in analyzing change processes so that they feel strengthened to sustain, extend and expand change, and advancing advocacy for gender equality and social justice (adapted from Batliwala and Pittman, 2010). 4

---

3. I am not going to discuss the differences of each terminology in this paper. According to Podems, gender approaches might identify or record the differences between men and women, while feminist evaluations would explore why these differences exist and “challenges women’s subordinate position”: Podem, D (2010), Feminist evaluation and gender approaches: There’s a difference? Journal of Multi-Disciplinary Evaluation, 6(14), page 8
4. Ibid, page 188
3.1 UN System Wide Action Plan (UN-SWAP) for gender equality and the empowerment of women

IN APRIL 2012, the UN System Wide Action Plan (UN-SWAP) was adopted at a meeting of the UN Chief Executives Board of Coordination to apply it throughout the UN system.

The UN-SWAP is:

- A unified gender equality framework, designed to promote accountability, a common understanding, enhanced coherence, systematic self-assessment, and steady targeted and progress approach to which the UN system entities can aspire and adhere in the work on gender equality and the empowerment of women at the corporate level. For the first time the UN system has a common set of measure on gender mainstreaming and the representation of women against which to report, and against which senior managers can be held accountable.

- The UN-SWAP will facilitate and substantiate accountability to Member States for the work of the UN system on gender equality and the empowerment of women.

- Its conceptual model involved progression from a checklist to scorecard approach, transforming accountability for the work of entities on gender equality from a reactive to a proactive one.

The UN-SWAP includes a set of 15 common system-wide performance indicators applicable to all UN entities, and UN entities will be expected to meet all of the requirements by reporting for 2017.

According to baseline findings from 2013 UN-SWAP reporting, the main findings are as follows:

- Considerable improvement is required; overall, only 31% of UN system “meets” or “exceeds requirements”, while 43% “approaches requirements”.

- The importance of a majority of entities rating themselves as “approaches requirements” should not be underestimated.

- The UN system as a whole exceeded minimum requirements in only 9% of individual ratings, demonstrating that while it evidences some good practice, there is still much improvement need to most fronts.

Based on the 2013 UN-SWAP findings, lessons learnt is summarized as below:

- Promotion of gender equality and the empowerment of women requires a conceptual model that defines clearly roles and responsibilities for all staff, built on inter-governmental mandates.

- Senior managers only support accountability measures if they are clear about resources and capacity needed to implement these measures.

- Effective coordination revolves around mutually defined goals, trust built by the coordinating body, and inclusive.

- To ensure that the UN system is more than the sum of its parts, a trusted credible and effective entity needs to play the coordinating role working synergistically with vital network and professionals from across entities and levels.
3.2 UN-SWAP Evaluation Performance Indicators

THE UNITED NATIONS Evaluation Group (UNEG), which brings together all the evaluation offices of the UN, has played an important role in setting the evaluation Norms and Standards to highlight the need for people centered evaluation.

In 2013, following the conceptual framework of the UN-SWAP, the UNEG developed UN-SWAP Evaluation Performance Indicator Technical Note. According to the Technical Note, integrating gender dimensions in evaluations is still a relatively new area of practice in evaluation. Institutional and methodological challenges exist and a shift in the way evaluations are conducted is required.

Furthermore, the Technical Note says: “the development and testing of new gender responsive approaches and methods must also be undertaken and fully implemented. Therefore, it is reasonable to expect that UN entities may not be in a position of “meeting requirements” for this indicator immediately, and 3-5 year period if more realistic.

According to the UNEG, as stated in the Technical Note, gender equality responsive evaluation is as one that incorporates the gender equality mainstreaming principles into evaluation such as quality, inclusion and non-discrimination. By doing so, such evaluation contribute to the social and economic change process that is central to most development programming by identifying and analyzing gender inequalities, discriminatory practices and unjust power relations that are central to development problems.

The Technical Note includes the UN-SWAP Evaluation Scorecards. It is a reporting tool organized around 13 scoring criteria which capture the overall elements related to mainstreaming gender equality throughout the evaluation process. The 13 criteria include: design, documents (evaluability assessments, ToR, inception reports), evaluation reports, management responses and evaluation guidance documents.

7. Ibid.
8. Ibid.
10. Ibid.
11. Ibid.
12. Ibid.
3.3 Integrating Human Rights and Gender Equality in Evaluation – Towards UNEG Guidance

IN THE UN, in response to a noted gap in evaluation guidance in general and the UN system-wide mandates to integrate Human Rights & Gender Equality (HR & GE) in all areas of work, including evaluation, the UNEG13 HR & GE Task Force14 reviewed existing frameworks and tools and developed the Integrating Human Rights and Gender Equality in Evaluation – Towards UNEG Guidance in 2011.15

By doing so, the UN system will be better able to learn lessons, hold key stakeholders accountable for results, and in turn improve policies and programming, which will contribute to the realization of HR & GE and meeting the Millennium Development Goals (MDGs) and other UN mandates.16

The Guidance is primarily tailored for use by UN evaluators conducting evaluations of programmes and projects, and is particularly useful for evaluation of interventions nor specifically focused on either human rights or gender equality. It can also be helpful to those designing and planning new interventions to support their analysis of HR &GE and promote their inclusion in the design of interventions and Monitoring and Evaluation (M&E) systems. With some adaptation by users, it can also be used as a tool to support other types of evaluation carried out within and outside the UN system.17

According to the HR & GE Guidance, principles of HR&GE evaluation are:

INCLUSION: Evaluating HR & GE requires paying attention to which groups benefit and which groups contribute to the intervention under review. Groups need to be disaggregated by relevant criteria: disadvantaged and advantaged groups depending on their gender or status (women/class, ethnicity, religion, age, location,etc.) duty-bearers of various types, and rights-holders of various types in order to assess whether benefits and contributes were fairly distributed by the intervention being evaluated.18

PARTICIPATION: Evaluating HR & GE must be participatory. Stakeholders of the intervention have a right to be consulted and participate in decisions about what will be evaluated and how the evaluation will be done. In addition, the evaluation will assess whether the stakeholders have been able to participate in the design, implementation and monitoring of the intervention. It is important to measure stakeholder group participation in the process as well as how they benefit from results.19

FAIR POWER RELATIONS: Both HR & GE seek, inter alia, to balance power relations between or within advantaged and disadvantaged groups. The nature of the relationship between implementers and stakeholders in an intervention can support or undermine this change. When evaluators assess the degree to which power relations changed as a result of an intervention, they must have a full understanding of the context, and conduct the evaluation in a way that supports the empowerment of disadvantaged groups.20

MIXED EVALUATION METHODS: regardless of the size of an intervention, evaluating HR & GE should use both quantitative and qualitative methods. The former can give credible information about the extent of results for particular groups of stakeholders, while the latter can assist in explaining how those results are achieved. Qualitative methods also allow for the voice of the most vulnerable to be heard. Information from mixed methods can assist in the triangulation of data, increasing reliability and validity, as well as being useful for exploring whether/why different stakeholders groups benefited differently.21

The HR & GE Guidance is organized to follow the main stages of evaluation planning and implementation:

1. Preparing for an evaluation: including how to assess the evaluability of the HR & GE dimensions in an intervention and how to deal with different evaluability scenarios.
2. Preparing the evaluation Terms of Reference (ToR): including step-by-step guidance on developing the elements of an evaluation ToR.22
3. Implementing the evaluation: including the importance of selecting and using appropriate methods of an evaluation to ensure that the HR & GE aspects of the intervention will be identified and analyzed during the evaluation process.23

The UNEG is a professional network that brings together the units responsible for evaluation in the UN system including the specialized agencies, funds, programmes and affiliated organizations. UNEG currently has 43 such members and observers. http://www.uneval.org/

9 UN agencies (PAO, ILO, OHCHR, UNICEF, UNDESA, UNDP, UNICEF, UNIOios, and UN Women) were the HR & GE Task Force members.

http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=1401


Ibid.

Some groups may be negatively affected by an intervention. An evaluation must acknowledge who these stakeholders are and how they are affected, and shed light on how to minimize the negative effects.
Challenges of Integrating Gender Equality in Evaluation

19. Ibid.
20. Ibid. e.g. women’s empowerment where women are disadvantaged gender within a given context. In addition, evaluators should be aware of their own position of power, which can influence the responses to queries through their interactions with stakeholders. There is a need to be sensitive to these dynamics.
21. Ibid. Finally, mixed methods can help to compensate for bias in privileging certain ways of knowing and communicating (e.g. literacy, ability to speak, publicly, conceptualizing in ways that resonate with the evaluator’s frame of reference).
22. Ibid., page 11 This includes how to conduct a stakeholder analysis, how to define evaluation criteria, evaluation questions and indicators and criteria for selecting an evaluation team that take HR & GE into consideration.
23. Ibid. This outlines the process of writing the evaluation report, dissemination of evaluation results and responding to evaluation findings and recommendations.
IN ORDER TO ENSURE gender equality in evaluation, we have been applying the above mentioned Technical Note and Guidance when managing evaluations in UN Women. Overall quality of the evaluations has been significantly improved due to the organization’s efforts and investment in the last years. Yet, based on our experience in managing evaluations in UN Women Asia and the Pacific region, degrees of gender equality in evaluation are still diversified.

In this section, we would like to examine some challenges of integrating gender equality in evaluation using real examples.

4.0 CHALLENGES OF INTEGRATING GENDER EQUALITY IN EVALUATION

4.1 Conceptual difficulties

AS MENTIONED ABOVE in 3.3, according to the UNEG, the principles of HR &GE evaluation are: Inclusion, Participation, Fair Power Relations and Mixed Methods. We would like to examine conceptual difficulties of applying the principle to evaluations.

According to the UNEG Guidance, gender equality in evaluation need to “disaggregate groups by relevant criteria”, and “whether the stakeholders have been able to participate in the design, implementation and monitoring of the intervention. It is important to measure stakeholder group participation in the process as well as how they benefit from results conduct the evaluation in a way that supports the empowerment of disadvantaged groups.

When applying the concept of “participation” to evaluation, we have encountered challenges such as:

- There is no agreed-upon method for the measurement of participation;
- Indicators of participation focused on participation in groups; and
- Little attention has been paid to qualitative indicators of participation.

In many gender focused programmes, an outcome is “women’s participation in a committee or a group” and the indicator for that is “percentage of women’s representation”. When conducting an evaluation, we often found out that women’s participation is not the right level of outcome (3-5 years’ achievement) and, since the indicator is quantitative, qualitative aspect of women’s participation is overlooked in monitoring.

For instance, UN Women Viet Nam Country Office’s (CO’s) Evaluability Assessment (EA) of the Programme on strengthening women’s capacity in disaster risk reduction to cope with climate change in Viet Nam (2013) argues:

“the representation of Women’s Union membership of Committees for Floods and Storms Control (CFSC) has been achieved in 2013. The activities under this programme have contributed to the success in the negotiation to include Women’s Union (WU) in CFSC. The next challenge that WU faces is how to bring in gender and women’s rights issues in the discussion in CFSC and to contribute to developing a gender-responsive Disaster Risk Reduction (DRR) plan.”

24. For instance, Report on the UN Women evaluation function (2012) rated 5 evaluations in Asia and the Pacific region either “excellent” or “good”.
25. Thus, the principle of “mixed-method” will be discussed later in methodological challenge.
26. Please see page 8.
27. UN Women Viet Nam Country Office (CO), Kusakabe.K, Evaluability Assessment (EA) of the Programme on strengthening women’s capacity in disaster risk reduction to cope with climate change in Viet Nam (2013), page 5 http://gate.unwomen.org/evaluationadmin/manageevaluation/viewevaluationdetail.html?evalid=4749
4.2 Methodological difficulties

THE UNEG HR & GE Guidance says that gender equality in evaluation should use both quantitative and qualitative methods.

Mixed-method combines the strengths of quantitative methods with the ability of qualitative methods to describe in depth the lived-through experience of individual subjects, groups or communities. These benefits are particularly important for gender equality in evaluation.

Recognizing the importance of mixed evaluation methods, we, as an evaluation manager, always mention “mixed methods” in the ToR for evaluation and the inception report. However, most of our programme evaluations heavily weighted on qualitative methods, and we analyzed the reasons for that as follows:

- Evaluator’s preference in qualitative research only;
- Evaluator’s lack of competencies in quantitative methods;
- Programme stakeholders’ reluctance in responding to quantitative data collection.

Based on our experience in managing gender-focused programme evaluations in Asia and the Pacific region, it can be said that there is a general tendency to weight more on qualitative methods rather than quantitative methods.

It is an assumption that one reason is feminism study is based on more qualitative research methods to seek to correlate social and cultural construction. Therefore, as a critical approach to the programme, the evaluators prefer to use qualitative methods and are thus shy away from quantitative and macro data.

Furthermore, due to the weight on qualitative research in education, our assumption is that some evaluators do not have competencies in applying quantitative methods to evaluation. Sometimes, we could observe wrong sample size and questionnaire design.

It is also a typical challenge that not many stakeholders respond to a survey, but they speak out when interviewing them individually. There are various factors for it, such as time constraints, sensitivity and power-relations. But stakeholder’s lack of understanding in evaluation mixed methods could be one of the reasons for the low survey response rate.

In many programmes, in our experiences, the concept of legal and political empowerment is relatively defined and the indicators are quantified, but measuring economic and social empowerment is often in its relative infancy.

Furthermore, applying the concept of “empowerment” is also another area of challenge, because:

- Empowerment is difficult to measure, and there is not an agreed-upon method for measuring it;
- The term “empowerment” is often not clearly defined, thus indicators are weak;
- Indicators of empowerment do not encompass both personal and socio-economic and political change.

In many programmes, in our experiences, the concept of legal and political empowerment is relatively defined and the indicators are quantified, but measuring economic and social empowerment is often in its relative infancy.

29. The UNEG Integrating Human Rights and Gender Equality in Evaluation – Towards UNEG Guidance actually uses the language empowerment in the Guidance. It says “there are also criteria that can be applied to evaluations that are derived directly from the HR & GE principles of equality, participation, social transformation, inclusiveness, empowerment, etc. and their use is strongly encouraged” (page 25). The criteria are indeed interrelated.

28. The UNEG Integrating Human Rights and Gender Equality in Evaluation – Towards UNEG Guidance actually uses the language empowerment in the Guidance. It says "there are also criteria that can be applied to evaluations that are derived directly from the HR & GE principles of equality, participation, social transformation, inclusiveness, empowerment, etc. and their use is strongly encouraged" (page 25). The criteria are indeed interrelated.

29. This includes permitting unbiased generalizations to the total population, precise estimates of the distribution of sample characteristics and breakdown into subgroups, and testing for statistically significant differences between groups. M. Bamberger and M. Segone, How to design and manage Equity-focused evaluations, UNICET, 2011, Page 51

30. This is the main reason why quantitative results are often not used in our evaluations. The response rate is not sufficient to come to a meaningful conclusion.
4.0

CHALLENGES OF INTEGRATING GENDER EQUALITY IN EVALUATION

4.3 Framework difficulties

IN THE FIELD of gender equality in evaluation, there is always a problem of creating a “hybrid” evaluation framework. There are three frameworks in general:

1. Causal framework
2. Contribution framework
3. Gender analysis framework

CAUSAL FRAMEWORK is to demonstrate the causal chains leading to programme impact. The Logical frameworks, Theory of Change (ToC), and Results-Based Management (RBM) are among the causal frameworks. This approach aims to systematize and identify logical hierarchy, and outlines how project objectives are reached. It involves a cause and effect analysis, a stakeholder analysis, and objectives tree and hierarchy, and an implementation strategy. Results, in this approach, may be understood as outputs, outcomes and impact.

The UN had adapted the RBM approach and, therefore, the UNEG HR & GE Guidance also recommends applying ToC to evaluation.

CONTRIBUTION FRAMEWORK is to track the multiple and variable forces involved in producing change, and highlight the contribution of change agents to the change process and intended outcomes. The prominent contribution frameworks are outcome mapping and participatory approaches. Outcome mapping tracks outcomes, resulting from changes in behavior, relationships, activities, or stakeholders. Typically, outcomes and progress markers are identified for each stakeholder on a three-point scale rating from “expect to see, like to see, and love to see”. The main aim of participatory approaches is to empower the various stakeholders – community members, project staff members and facilitators – to initiate, control and take corrective action themselves.

This UNEG HR & GE Guidance does not refer to this framework. Yet, it is a well-known approach and, since the one of the key principles of HR & GE evaluation is participation, we as a gender organization particularly pay attention to participation process and methodology in evaluation.

GENDER ANALYSIS FRAMEWORKS are also classified as alternative frameworks in evaluation, however, it is the most commonly used by gender evaluators with feminist inclinations. The Harvard analytical framework, the Moser gender planning framework and social relations approach are among the prominent gender analysis framework. A summary of the frameworks are as follows:

- HARVARD ANALYTICAL FRAMEWORK: identify the different resources and constraints facing women and men in social and economic development, and examine the implications of women’s multiple social and economic roles.
- MOSER GENDER PLANNING FRAMEWORK: address women’s “three roles” in production, reproduction, and community management, and the implication these roles have for women’s participation in the development process.
- SOCIAL RELATIONS APPROACH: locate the family and household within the network of social relations, connecting them to the community, market and states.

The UNEG HR & GE Guidance only refers to “HR & GE analysis”, but in gender field, they are the most prominent frameworks.

Applying the HR & GE principles means that we apply the above mentioned frameworks to evaluation and it is technically challenging to create a “hybrid” model in our daily work. What often happens is that evaluation framework becomes “a bit of everything framework” – mostly relying on ToC with some elements of contribution and gender analysis. But the problem of this “hybrid” approach is that evaluation framework becomes unclear. It is, in fact, seldom to identify evaluators who have knowledge, skills and experiences of applying all frameworks in evaluation. Also, in our viewpoint, gender analysis frameworks are applied to programme planning rather than evaluation, and application of the frameworks to evaluation may require more practical guidance.

Another challenge of applying the frameworks to gender equality evaluation is that the frameworks are bound by six-month to three-year grant periods whereas very little can be realistically measured in such short periods of time as the pace of long-lasting change is usually slow. This criticism may not be unusual to evaluations of other social development programmes. However, in the context of gender equality evaluation, it is particularly an issue, as change in gender norms or practices require a long period of time.
Measuring a relatively short-term result also fails to assess a conflict or a backlash which may occur in a long time even after completing the programme. A common instance of this in Asia and the Pacific is that “domestic violence against women by men rise immediately after the setting up of women’s microcredit groups as men feel threatened by the possible economic independence that women could derive from their association with these groups.40 In most cases, violence declines over time when the groups’ benefits not only to women but to the entire household become apparent to men. But a conventional evaluation framework is not likely to capture the initial conflict/backlash, and if it does, quite likely it will be interpreted as a sign of “failure” of the programme intervention.41

In addition, new debates on evaluation design, experimental (randomized) and quasi-experimental poses another challenge on gender equality in evaluation. It is technically challenging to identify and collect data from control groups of sensitive issues such as GBV. Some of these designs do not address the perspectives and experiences of marginalized groups in the evaluation process.42 Gender equality in evaluation (feminism evaluation, empowerment, participatory and transformative approaches) starts with the principle of including those voices.

37. Ibid.
38. Ibid. It is more broadly oriented than the other frameworks and draws on explicitly feminist root.
39. Ibid.
40. Association for Women’s Rights Development (AWID) http://www.awid.org/
41. R. Govinda, Mapping “Gender Evaluation in South Asia”, Indian Journal Of Gender Studies, Volume 19, Number 2, June 2012, page 197
42. K. Hay, engendering Policies and Programmes through Feminist Evaluation, Volume 19, Number 2, June 2012, page 329
5.1 Strengthen gender sensitive indicators

DEMONSTRATING RESULTS WILL continue to be a key issue in international development, and those results can only be shown quantitatively and qualitatively. Thus, in order to strengthen gender equality in evaluation, UN Women has been strengthening programme indicators.

The best place to strengthen programme indicators is in its planning stage. When developing indicators, first, we review elements of good indicators with programme managers. In the RBM, SMART indicator is well-known as:

- **Specific (Who and What?)**
- **Measurable (How much?)**
- **Area specific (Where?)**
- **Relevant and/or Realistic**
- **Time-bound (When?)**

Following the concept of SMART, we further define characteristics of a good indicator as follows:

- **INCLUDES A UNIT OF MEASUREMENT**: Does it specify a unit of measure? (e.g. Quantitative: #, %, rate, score, etc.; Qualitative: perspectives, types etc)
- **NON-DIRECTIONAL**: Does it avoid direction of change (“increase”)?
- **RELEVANT**: Does it measure the most important results you want to achieve among intended beneficiary group (institution, police, etc.)? It needs to be relevant to programme and national standards.
- **SPECIFIC**: Are all terms clear and can they be clearly defined?
- **MEASURABLE**: Do you have resources/methods to measure it?
- **VALID**: Does it really measure what you think it does?
- **RELIABLE**: Can it measure the same thing consistently over time?
- **SENSIBLE TIMEFRAMES AND FREQUENCY**: Can it be measured at time intervals that allow you to measure change (e.g. before/after project)?
- **BOTH QUALITATIVE AND QUANTITATIVE**
- **EASY TO INTERPRET AND ENABLE TO TRACK OF CHANGES OVER TIME.**

Then, we develop indicators in the following process:

- Develop in participatory manner
- Brainstorm all possible indicators and then select best ones
- Limit the number of indicators (2 to 3 per result)
- Examine the relevance to needs of the user
- Measure realization and enjoyment of rights
- Always sex-disaggregate

In the planning exercise, it is quite difficult to identify suitable indicators to meet the above criteria in the beginning. However, discussing in a group and examining the indicators according to the criteria certainly strengthen the quality of the indicators.

When the concept, such as participation and empowerment, does not have agreed-upon standard indicators, we examine the definitions of the concept according to international standards and guidance, and agree on the concept of the terminology. This helps us to breakdown components of the concept.

For instance, if we agree to the definitions of empowerment as:

“Empowerment is about-people - both women and men – taking control over their lives: setting their own agenda, gaining skills, increasing self-confidence, solving problems, and developing self-reliance. It is both a process and an outcome”, and “A term generally used to describe a process by which powerless people become conscious of their own situation and organize collectively to gain greater access to public services or to the benefits of economic growth”.43

---

43. CIDA (1997), Guide to Gender Sensitive Indicators, page 39. These definitions are from CIDA.
44. Ibid., page 40.
45. Ibid.
46. Ibid.
5.2 Planning evaluations in programme cycle

GENDER SENSITIVE PLANNING is a key to conduct a gender equality evaluation. The UNEG HR & GE Guidance also says “it is much easier to evaluate the HR & GE dimensions of an intervention when they have been addressed during the design, implementation and monitoring and intervention.”

Therefore, UN Women Asia and the Pacific region has been trying to develop a M&E plan when developing new programmes. These programmes are normally 3-5 years programme, and in addition to conventional planning tools such as Logframe and Project Monitoring Framework (PMF), we develop a M&E plan, which includes sufficient resource allocation to M&E activities, and timing and types of evaluations. In many programmes, evaluation tends to be “after-thoughts”. When a donor requests to conduct an evaluation, it is often a final programme evaluation, but without adequate budget, ToC, and baseline data. Therefore, learning from our experience, we have been incorporating evaluation planning in programme planning, so that we can better ensure gender equality in the evaluation throughout the cycle.

For instance, UN Women Pacific Multi-Country Office (MCO) developed a M&E plan for the Pacific Regional Elimination Violence Against Women (EVAW) Facility Fund Programme for 2013-2017. The M&E plan includes an Evaluability Assessment (EA), Formative evaluation, Mid-term evaluation and Final evaluation. By doing so, the programme has been increasing its evaluability as well as reflexivity.

Planning evaluations in programme cycle resulted in producing more formative types of evaluation in the region such as EA in the past years, which had never occurred when the evaluation was “after-thoughts”. The formative evaluation often pointed out the weakness of ToC of our programmes, due to inappropriate application or lack of gender analysis. Without such analysis, the ToC often failed to identify what is actually the gender issue.

The advantage of identifying the problem at the early stage of the programme implementation through formative evaluations is that we can make an adjustment. For instance, we include “conduct a gender analysis to specify the gender issue” with a specific timeframe in the evaluation management response. Although this should have been done at the early stage of programme formulation, it is far much better than finding it out at in the final evaluation.

Planning evaluations in programme cycle also assists to design evaluations. It allows us to examine what realistic evaluation could be – whether it should be post-test only no comparison evaluation or pre- and post-test with no comparison group or quasi-experimental or experimental design. Depending on the type of the evaluation we choose, preparations for baseline data collection will differ.

Based on the components, two main areas to focus could be: (1) a personal change in consciousness involving a movement towards control, self-confidence and the right to make decisions and determine choices; and (2) organization aimed at social and political change.

The selected indicators should fit these components, but they will differ depending on the programme context and objectives. Therefore, it is always important to go back to the characteristics of a good indicator. By doing so, types of empowerment, programme time-scale to achieve the objective, number of affected people and the degree of affection will be clarified.

In fact, this is one of the most difficult areas in gender equality in evaluation. Even we take the above mentioned process, it still remains as a complex problem, because it is the area of measuring individual state of mind, which is closer to the field of psychology, and researching on decision-making process such as household level requires time and in-depth qualitative research.
HOW UN WOMEN IS OVERCOMING THE CHALLENGES

5.3 Methodological improvement

TO MAXIMIZE GENDER sensitivity of evaluation methodology, UN Women has been taking the following measures:

GENDER SENSITIVE EVALUATION TEAM:
In order to capture gender and social inequality, discrimination, exclusion and abuse, it is imperative that the evaluators themselves are sensitive enough to capture these. Gender prejudice run so deep that many of us even do not notice discrimination. It is indeed not easy to examine evaluators’ sensitivity only on CV, and we thus examine it in interviews. It is important to have a mix of men and women evaluators. From our experience, an all-women or all-men team could not capture the nuances. In fact, individual methods are not “feminist” or “non-feminist”. The key point of enhancing gender equality in evaluation is the evaluation team’s ability to interpret valid and reliable data to speak the nature and change around the gender inequality the programme is attempting to address.

GENDER SENSITIVE EVALUATION QUESTIONS AND TRIANGULATION:
In UN Women, the standard evaluation management process is that we, as an evaluation manager, draft an evaluation ToR, and the evaluation team develops an inception report. Previously, in many cases, UN Women left the evaluators to decide evaluation questions and methodologies. Thus, time for reviewing an inception report was very short. In the last years, we have been changing this practice that we initially develop a list of evaluation questions and methodologies, particularly paying an attention to gender sensitivity and equality, in the ToR. Then, we ask the evaluation team to further refine the questions and methodologies in an evaluation framework in the inception report, so that we can ensure whether the evaluation is going to address gender sensitive questions with right methodologies. When examining the evaluation framework in the inception report, we are also triangulating methodologies, particularly balancing quantitative and qualitative methods.

STRENGTHEN QUANTITATIVE METHODS:
As mentioned in the previous pages, many gender evaluators with feminism background are strong at qualitative methods, and our evaluations are mostly rely on qualitative methods. But in order to enhance reliability and validity, applying mixed-methods are important in gender equality evaluation, and we thus have been trying to improve the quantitative methods in particular. Recently, UN Women Cambodia CO has conducted a self-evaluation on 2nd National Action Plan on Violence Against Women (NAPVAW) project. In the evaluation, they prepared 10 qualitative and quantitative questions and collected data from individuals in one-on-one semi-structured interviews. For instance, to examine an output (Increase capacities of stakeholders from government and civil society to formulate or influence national policies in line with international standards), the Office defined it as an effectiveness question and examined “percentage of government and civil society participants reporting significantly improved” “slightly improved” or “no change”. At the same time in the interview, they asked the partners to describe their capacities (qualitative). As a result, the evaluation concluded that 85% of respondents reported improved capacities to formulate or influence government policies (quantitative). Furthermore, representative quotes provide in-depth analysis which is not provided by the numbers.
5.4 Evaluation capacity development

IN THE PAST YEARS, UN Women has been intensively training its staff and partners on evaluation. For instance, in Asia and the Pacific region, 393 staff and partners were trained only in 2012. We would like to share our experience of building evaluation capacity in the region.

STAFF CAPACITY DEVELOPMENT: When conducting evaluation training for staff, we include a specific topic of evaluation which the Office currently needs. For instance, in case of evaluation training for the Viet Nam Office staff in August 2013, we identified that the Office needs support to conduct the above mentioned EA on the climate change project. Therefore, in the training curriculum, we included a section of what EA is, and drafted a ToR for the EA in the training workshop. In this manner, staff could immediately apply their knowledge on evaluation they have gained in the training to manage the actual evaluation. We have been taking this approach – identify the most needed evaluation skills and apply immediately after the training. Due to intensive advocacy on the RBM in the UN, most staff have a conceptual understanding of ToC, indicators and evaluation. However, the weakness is application. Therefore, as for staff training, we have been emphasizing on daily and on-site coaching by using the real evaluation example, i.e. “learning by doing”.

PARTNERS’ CAPACITY DEVELOPMENT: UN Women’s programme partners are mostly government officials and civil societies. We often conduct evaluation training exclusively for partners, but sometimes conduct it together with UN Women staff. According to our pre-and post training survey, partners’ average immediate knowledge gain after the training is not high (around 10-20% gain), due to diversified background and knowledge of the partners. Especially, in case training is conducted in English and is translated into a local language, the result of the average knowledge gain could be even lower. There is always a dilemma in partners’ evaluation capacity building – it is easier to conduct evaluation training for partners who speak English, but they may not be the group who need evaluation training most. In fact, in regard to partners’ capacity building, we have more challenges than solutions. One thing we have been trying is to include evaluation capacity building strategy in the above mentioned programme M&E plan. It is much easier to identify a group of partners under the programme and is also possible to work with the same partners over the period of 3-5 years. For instance, we have conducted evaluation training for partners in EVAW programme in Cambodia in 2013 and the plan is to measure the same group members’ knowledge and skills over the programme period (5 years).

EVALUATOR’S CAPACITY DEVELOPMENT: Evaluation is an evolving area in Asia and the Pacific region. Based on our experience in managing gender equality evaluations in the region, there are some prominent gender sensitive evaluators, but overall, the numbers of these people are still limited. In order to increase the number of such evaluators and to support strengthening their evaluation skills, for instance, UN Women provided bursaries to researchers and practitioners who are working in the field of gender and evaluation to present their research topics in national and regional evaluation conferences. After presenting the research topics in the conference, the researchers/practitioners develop an article on their topics and we conduct a peer-review among the group and with UN Women staff. UN Women also support to publish and disseminate the articles. Recognizing capacity building requires a long time network, we keep the relationship with the researchers and practitioners and provide them with an opportunity to practice their gender sensitive methods in evaluating UN Women programmes, and also to present their research findings among M&E networks.
This paper examined the latest development of the UNEG’s work on gender equality dimension in evaluations, which is influencing concepts, methods and frameworks of evaluation, particularly in the UN. Gender equality in evaluation is still a new developing area and, in fact, there are many challenges in practical work. Yet, it is important to analyze the difficulties and limitations, and demonstrate the strengths, rigor, and validity of the gender-sensitive approaches we are using. Particularly, in the circumstance where other evaluation approaches dominate evaluation discourse and many donors consider them as the credible approach, such work is essential.

In the end, examining gender equality in evaluation is about reflection, analyzing how societies are changing and which policies and programmes worked and which did not work, and how they could be refined for shifting norms and standards towards more equitable societies.
References

**Guidance**

UN Women (2013), System-Wide Action Plan for implementation of the CEB Policy on gender equality and the empowerment of women

UN Women (2012), UN-SWAP A plan to improve gender equality and the empowerment of women across the UN system

UN Women (2012), Frequently Asked Questions, UN System-wide Action Plan (UN-SWAP)


UN Women (2012), Evaluation policy of the United Nations Entity for Gender Equality and the Empowerment of Women


Canadian International Development Agency (CIDA) (1997), Guide To Gender-Sensitive Indicators

**Evaluation Report**

UN Women Viet Nam Country Office (2013), Kusakabe.K, Evaluability Assessment (EA) of the Programme on strengthening women’s capacity in disaster risk reduction to cope with climate change in Viet Nam

UN Women Cambodia Country Office (2013), Self-evaluation of AAP Project Phase 1, Outcome 1: 2nd NAPVAW Formulation (DRAFT)

**Programme Document**


**Journals**

Govinda, R (2012), Mapping ‘Gender Evaluation’ in South Asia, Indian Journal of Gender Studies, Volume 19, Number 2, 187-209


Bamberger, M.and Segone, M. (2011), How to design and manage equity-focused evaluations, UNICEF