Evaluation requires careful preparation to ensure it will be of high quality, credible and useful. This chapter discusses the evaluability assessment, stakeholder analysis and engagement, development of the ToR, and selection of the evaluation consultant(s) and/or firm.
An overview of the preparation stage is depicted in Figure 5. As a first step, an evaluation manager should be identified from within the respective UN Women office to manage the evaluation process. To maximize impartiality, the programme officer responsible for the implementation of the programme to be evaluated should not have individual decision-making responsibility in the evaluation process.

**Figure 5. Overview of the preparation stage**

1. **Appoint Evaluation Manager**
2. **Evaluability assessment**
3. **Stakeholder analysis and engagement**
4. **Develop ToR**
5. **Select evaluation team**

**A. CHECKING EVALUABILITY**

Ideally all programmes and projects should be designed to enable robust M&E. An evaluability assessment is a systematic process intended to determine whether or not an intervention is: in a condition to be evaluated, justified, feasible and likely to provide useful information. An evaluability assessment also helps prepare the programme to create the conditions necessary for an evaluation. It typically takes place at the beginning of an intervention but can be carried out at any point of the implementation. An evaluability assessment is not a replacement for a high-quality programme design. An evaluability assessment is useful for ensuring the intervention is ready for an evaluation. It is in the best interest of UN Women to either invest the time and funds necessary to hire external consultants to conduct an evaluability assessment or to have one conducted internally by the respective office (see Tool5. How to conduct an evaluability assessment).

The evaluability assessment will review:

- Programme design
- Availability of relevant monitoring information and data
- Conduciveness of the context for evaluation

---


17 According to the glossary of terms in “Evaluation and results-based management (RBM)” by the Development Assistance Committee, evaluability is the “extent to which an activity or a programme can be evaluated in a reliable and credible fashion”, and evaluability assessment calls for “the early review of a proposed activity in order to ascertain whether its objectives are adequately defined and its results verifiable.”
Box 4.

EVALUABILITY ASSESSMENT:
Viet Nam Country Office

Strengthening women’s capacity in disaster risk reduction to cope with climate change in Viet Nam.

The evaluability assessment was a chance for the country office to meet with project partners and beneficiaries to assess their understanding about the project. Their comments and feedback enabled the office to make necessary changes to the project design and implementation that would ensure the project could achieve its results, is measurable and would support a future evaluation. Implementing partners were closely involved in the evaluability assessment process to facilitate a sense of ownership in order to have a shared sense of responsibility and commitment to work on its recommendations together. Because evaluability assessments were not commonly undertaken in this context, the office had to spend time explaining to its partners the purpose of the exercise and that it was not an evaluation. This also enabled partners to provide constructive and honest feedback. It is important to know that an office conducting an evaluability assessment should take into consideration the time and efforts necessary to respond to recommendations and make improvements to the project after the evaluability assessment, which could result in delays in project implementation.

Vu Phuong Ly
Programme Specialist, UN Women Viet Nam
If the evaluability assessment determined that the programme can be evaluated or identified specific measures that can be taken by an evaluation in order to address programme shortcomings, the evaluation manager can move onto the next step: identifying stakeholders to be engaged in the evaluation process. Programme managers can begin making plans to implement the necessary changes. The suggested evaluation approach and methods should be assessed and incorporated into the ToR for the evaluation. The regional coordination and planning specialist can provide technical guidance and support on programme design aspects while the regional evaluation specialist can be consulted with regards to how to prepare for an evaluation.

B. STAKEHOLDER ANALYSIS AND ENGAGEMENT

A defining characteristic of gender-responsive evaluations is including stakeholders—including both women and men and vulnerable groups, such as indigenous peoples, persons with disabilities, etc. Below are several key principles to follow when engaging them.18

- **Inclusiveness:** Take into account all directly and indirectly affected groups—stakeholders, duty-bearers and rights-holders—and be sensitive to differences among them. Disaggregate groups by relevant criteria (sex, class, age, ethnicity, religion, etc.) and pay attention to which groups benefit and which groups contribute to the intervention under review.

- **Participatory and reflective:** Engage stakeholders in an active and meaningful involvement in the design, management and conduct of the evaluation. Assess whether or not the stakeholders were able to participate in the design, implementation and monitoring of the intervention under review and if the intervention (and the evaluation) reflects their engagement. While stakeholders will need to be contacted (e.g., through interviews or surveys) during the evaluation to provide information to the evaluators, gender-responsive evaluation goes beyond this by ensuring active participation in or co-ownership of the evaluation process.

- **Respect:** Treat all stakeholders, particularly those who are vulnerable, with respect for their culture, language, gender, location, and abilities, and develop appropriate ways to engage and be accountable to them.

---

18 The term **stakeholder** is broadly used to include those who deliver, influence and are impacted by the programme. Evaluation stakeholders are people who have a vested interest in evaluation findings.
• **Transparency and accountability:** Ensure the design and conduct of the evaluation is transparent and responsive to questions about all aspects of the process. The results should be publicly accessible (in languages and other formats that stakeholders can access), and feedback should be provided to stakeholders about the process, results and use of the evaluation.

**Benefits of stakeholder participation in gender-responsive evaluation**

Thinking about how to identify and engage women and men, as well as vulnerable groups, duty-bearers, rights-holders, and other stakeholders at the outset of an evaluation has several benefits and is instrumental to building ownership, promoting accountability and encouraging evaluation use.

• **Contributions:** Different perspectives will enrich the evaluation design and approach, including: identifying gender equality and human rights issues and implications that may not have been explicit in the programme design and implementation; identifying and collectively finding solutions for evaluation constraints and challenges; facilitating and guiding the conduct of the evaluation; and bringing realistic, on-the-ground insights to the review and assessment of evaluation findings, conclusions and recommendations.

• **Use and usefulness:** The ultimate definition of evaluation success is the extent to which it is used and useful. Participatory approaches to evaluation have demonstrated increased credibility of evaluation results and thus use. Early engagement of stakeholders manages expectations for what the evaluation can provide by involving key players in the definition of why it is being done. Stakeholder engagement is instrumental to building ownership and promoting evaluation use.

• **Capacity building:** Being involved in the evaluation process is itself a learning experience and can serve to build the capacity of stakeholders through increased exposure to gender equality and human rights issues and gender-responsive evaluation approaches. It can help draw explicit connections between programmes and the larger objectives of social change; encourage good practice in tracking and measuring gender equality and human rights; and can be one step in helping duty-bearers (i.e., government officials or authorities) to become committed to gender equality and human rights obligations.

• **Accountability:** Bringing together duty-bearers and rights-holders generates a space for mutual accountability, transparency, and application of key gender equality and human rights principles.

• **Empowerment:** Engaging stakeholders and beneficiaries in all stages of an evaluation process and providing a space for them to determine how a meaningful
process of reflection and assessment should be undertaken can empower participants to take ownership of development interventions.

Identifying stakeholders

The evaluation manager should undertake a stakeholder analysis using the stakeholders analysis matrix (see Tool 9. Stakeholder analysis template) to identify who should participate, how and when, as well as the relevance of a particular group to the purpose and use of the evaluation (see Box 5). While there are many benefits to involving stakeholders in the design, planning and conduct of the evaluation, the evaluation manager will need to weigh these benefits against the desired level of impartiality required of the evaluation. Additionally, the participation of all stakeholders in the evaluation process can be challenging to manage, and may have cost and time implications.

Box 5.
Key questions for identifying stakeholders

Who: Stakeholders, disaggregated as appropriate
What: Their role in the intervention
Why: Gains from involvement in the evaluation
Priority: Importance of involvement in the evaluation process
When: Stage of the evaluation to engage them
How: Ways and capacities in which stakeholders will participate

Once the universe of stakeholders is identified, it is important to define a sub-set of actual evaluation users, narrowing the list of potential stakeholders to a much shorter, more specific group of primary intended users. As mentioned, there is a need to carefully balance the desire to be inclusive (to maximize broad input) against the challenge of managing the evaluation process efficiently.

Engaging stakeholders

Evaluation processes should clearly define an organization and management structure and establish the roles and responsibilities of the key stakeholders. Key stakeholders can be involved in the evaluation process through the establishment of the management group and the reference group.
Establish an evaluation management group

An evaluation management group should be established to oversee the evaluation process and should be coordinated by the evaluation manager. This group should consist of UN Women staff and, in the case of joint evaluations, may also include non-UN Women staff. The responsibility for final approval of the evaluation ToR, selection of the external evaluation team, inception report and final evaluation report should be with the evaluation management group. The group should comprise members from senior management, M&E officers or focal points, and the programme officer responsible for the programme that is to be evaluated. In the case of joint evaluations, representatives of the partner entities may also be included. The programme officer should ensure that the evaluation addresses the information gaps and evaluation questions relevant to the programme being evaluated. Regional evaluation specialists could also be considered as members of the evaluation management group for evaluations managed by the RO, MCO or CO in an advisory capacity. The evaluation management group should be chaired by the representative or by a member of the senior management. A ToR for the management group that outlines their main roles and responsibilities should be developed (see Tool 11. Management group terms of reference template). Involvement in the evaluation management group can be light (i.e., via e-mail), in order to lessen the administrative burden. The main goal is to ensure senior management oversight of all evaluation processes.

Establish a reference group

The evaluation reference group is an effective way to engage stakeholders, as it provides for their systematic involvement in the evaluation process. Careful selection of participants and clear definition of their roles and responsibilities is important for ensuring the best use of a reference group.

The evaluation manager can use the stakeholder analysis to aid in the selection of the key stakeholders for the reference group. Their role(s) can vary, for example to include decision-making, providing contextual or technical expertise and/or supporting evaluation dissemination and use. The evaluation manager needs to ensure that there is a clear understanding amongst the reference group members on how they will contribute to the evaluation process, which can be done through the establishment of a ToR that outlines the main roles and responsibilities of the reference group. Limiting the number of participants may facilitate efficient management.
The main functions of the reference group include (See Tool 12. Reference group terms of reference template):

- Facilitating the **participation of the key stakeholders** in the evaluation design, defining the objectives, the evaluation scope and the different information needs.
- Providing input on the **evaluation products**: a) ToR, which defines the nature and scope of the evaluation; b) inception report, which defines the approach and methodology of the evaluation team; c) preliminary findings, which identify the key findings from preliminary analysis; and d) draft and final reports, to identify factual accuracy, errors of interpretation or omission of information.
- Providing **relevant information** (i.e., via surveys, interviews, etc.) and documentation to the evaluation team.
- **Disseminating evaluation results**.
- **Implementing evaluation recommendations** as appropriate (see Chapter 7).

### Box 6.
**Steps for managing the reference group**

1) Identify key stakeholders for the reference group as early as possible in the evaluation process.
2) Develop a ToR for the reference group to ensure clarity regarding member roles and responsibilities (Tool 12. Reference group terms of reference template).
3) Convene the reference group to discuss feedback on the evaluation process and the ToR for the evaluation.
4) Keep reference group members informed via e-mail or conference call (as necessary) as the evaluation proceeds.
5) Convene the reference group for the presentation on the preliminary findings of the evaluation.
6) Maintain an audit trail of comments on the evaluation products so that there is transparency in how the evaluation team is responding to the comments (see Tool 7. Evaluation product comment template).
Box 7.

ENGAGING WITH THE EVALUATION REFERENCE GROUP:
Regional Office for Asia and the Pacific

*Evaluation of the project on Regional Mechanisms to Protect the Human Rights of Women and Girls in South East Asia.*

The project, which involved two regional human rights bodies from the Association of South East Asian Nations, which covers 10 countries in South East Asia, established a steering committee (comprising of UN Women management and the donor) from its inception in 2010 that functioned as an evaluation management group for the final evaluation. Following the inception meeting of the evaluation process, UN Women Regional Office for Asia and the Pacific set up a stakeholder reference group to oversee the planning and progress of the evaluation. However, because the key partners were intergovernmental bodies, it was difficult to limit the number of stakeholders in the reference group because of the politically sensitive nature of relationships.

Therefore, to review the report findings and recommendations, the Regional Office for Asia and the Pacific broadened the group to ensure comprehensive stakeholder participation and a participatory approach. All concerned stakeholders provided their feedback to the report and its recommendations, ranking them as high, medium and low priority. These recommendations and the discussions from this meeting were inputs into the next phase of the programme.

This participatory approach facilitated buy-in, quality control, management of expectations, disclosure of findings, sharing of recommendations, and ownership in implementing the recommendations. This evaluation process—wherein an independent group of people spoke to the stakeholders, shared their findings with them, and involved them in designing the next phase of the project—was important in cementing the relationship of trust built by UN Women with the Association of South East Asian Nations and it is important in the context of the new UN Women- Association of South East Asian Nations Memorandum of Understanding.

Deepa Bharathi,
*Regional Programme Manager (CEDAW South East Asia Programme)*

Yumiko Kanemitsu,
*Regional Evaluation Specialist*
UN Women Regional Office for Asia and the Pacific
Box 8.
SIX COMMON CHALLENGES AND TIPS FOR MANAGING REFERENCE GROUPS

**CHALLENGE 1:** Reference group members may not be familiar with evaluation.
**TIP:** Spend time in the beginning to educate them on what an evaluation is, why it is done and the process.

**CHALLENGE 2:** Reference group members are not clear about their role in the evaluation and their responsibility to provide feedback.
**TIP:** Develop a clearly defined and agreed-upon ToR for the reference group at the beginning of the evaluation process that defines roles and manages expectations.

**CHALLENGE 3:** Reference group members do not have time to participate in meetings and/or provide feedback to evaluation products as agreed.
**TIP:** Allocate adequate time for feedback and alert them about shifting deadlines in advance. Identify alternative means of communication in order to enable stakeholders to meaningfully participate.

**CHALLENGE 4:** Reference group members can feel disconnected and become disengaged with the evaluation process.
**TIP:** Make sure that initial communications are set up well and there is a two-way communication. Consider holding regular workshops/teleconferences to facilitate communication and make members feel more connected.

**CHALLENGE 5:** Reference group members may change their perspective during the course of the evaluation on what should be included in the scope of the evaluation.
**TIP:** Ensure that the reference group is clear that due to contractual purposes and to ensure validity, the evaluation team must not alter the scope of the evaluation agreed upon in the ToR and clarified in the inception report.

**CHALLENGE 6:** All reference group members’ comments and feedback may not be reflected in the final report.
**TIP:** It is important for the evaluation report to reflect the diversity of views gathered from different stakeholders (programme managers, donors, beneficiaries, etc.), but at the same time evaluators must maintain their impartial judgement in the final report. Use Tool 7. Evaluation product comment template to track comments provided by the reference group and response from the evaluation team. Also provide them with a copy of UN Women Evaluation Policy and UNEG Norms and Standards and reiterate the need to ensure the impartiality of the evaluation, its rigor and integrity.
The participation of stakeholders in the evaluation process can be challenging to manage and may have cost and time implications (see Box 8). However, it is instrumental to building trust and accountability, and ensuring the relevance, quality, learning from and use of the evaluation. It is also a means for building national ownership and capacity in gender-responsive evaluation techniques. Evaluation managers need to pay particular attention to stakeholders who are very relevant to the evaluation, but who are difficult to engage due to various constraints, and find ways to ensure their participation while upholding ethical principles for engagement of stakeholders (see below section on ethical considerations for evaluation methods). For example, means of sharing information, choice of meeting location, timing and language used by the reference group may all have a bearing on the capacity of particular members to participate (e.g., rural or indigenous women).

C. DEVELOPING AN EVALUATION TERMS OF REFERENCE

The evaluation ToR is a critically important document in preparing for an evaluation. The ToR defines why the evaluation is being undertaken (purpose and objectives), what it will examine (scope), how (design and methods), when it will be conducted (time frame), who will use it (intended users) and how it will be used when completed. This section discusses the overall content and provides suggestions for the development process (see Box 9 and Tool 6). Subsequent sub-sections elaborate on how to define the evaluation questions, scope, approach and methods. The time and effort spent in preparing a good ToR has big returns in terms of quality, relevance and usefulness of the evaluation to be conducted.

Key components of the evaluation ToR

The background section should include an overview of the programme or project context and rationale for evaluation. It should also refer to the guiding documents for evaluation at UN Women, including the Evaluation Policy, Evaluation Chapter of the POM, the GERAAS evaluation report quality checklist, the United Nations System Wide Action Plan Evaluation Performance Indicator (UN-SWAP EPI) and this Evaluation Handbook. These documents serve as the frame of reference for the Evaluation Manager and the evaluation consultant(s) for ensuring compliance with the various requirements and assuring the quality of the evaluation report.

The description of the programme/intervention should include: a detailed outline of the programme (components, implementation status, key participants, budget),
and the logic and underlying assumptions upon which the strategy was developed (the theory of change). Any major divergences from the original programme strategy should be stated and explained. The resources and management structure of the programme should be described.

**Box 9. Outline of an evaluation terms of reference**

I. Background (programme/project context)
II. Description of the programme/project
III. Purpose (and use of the evaluation)
IV. Objectives (evaluation criteria and key questions)
V. Scope of the evaluation
VI. Evaluation design (process and methods)
VII. Stakeholder participation
VIII. Time frame
IX. Expected deliverables (including standard table of contents for an evaluation report; see Chapter 6)
X. Management of evaluation
XI. Evaluation team composition, skills and experiences
XII. Ethical code of conduct

Annex 1 UN Women GERAAS evaluation quality assessment checklist
Annex 2 UN Women Evaluation Consultants Agreement Form
Annex 3 UNEG Norms and Standards for evaluation
Annex 4 UN Women Evaluation Handbook

The description of the purpose of the evaluation should include: the why (what triggered it), how the evaluation results will be used and by whom. The purpose of the evaluation could be for learning, accountability or decision-making (see Chapter 2).

**Defining evaluation objectives**

The evaluation objectives should follow the purpose. They should be limited in number and clearly formulated considering the programme information available, the context in which the programme is being implemented, and the context in which the evaluation will take place. The objectives should be framed from a gender equality and human rights perspective. The objectives often identify the evaluation criteria upon which the programme or intervention will be assessed.
Box 10. INVESTING TIME IN DEVELOPING THE TERMS OF REFERENCE: Multi-country Office Kazakhstan

MCO Kazakhstan managed three decentralized evaluations in 2012-2013:

1. Final evaluation of women connect across conflicts: Building accountability for implementation of the UN SCR 1325, 1820, 1888, 1889 cross regional project
2. Final evaluation of women for equality peace and development in Georgia project
3. Mid-term evaluation of women’s economic empowerment project in Moldova

Much consideration was put in designing the ToR for the evaluations, with special attention paid to the evaluation criteria, evaluation matrix and expected usage and application of RBM. The initial draft of the ToR was prepared by the M&E focal point. Then the programme officer in charge of the thematic portfolio, project managers and the IEO reviewed it prior to sharing with the broader reference group. As an RBM and learning-based organization, a special emphasis in developing evaluation criteria was put on achievement of results, relevance of programme interventions to country development context and recommendations to UN Women on programming. Due to low evaluation capacities available in the region, evaluation ToRs anticipated involvement of a senior international evaluator as a team leader and a national evaluator/assistant evaluator to enable capacity development in evaluation for local consultants. A reference group meeting was convened to validate the ToR, in particular the evaluation scope, criteria and intended usage. The final draft was then shared with our respective geographic section at UN Women HQ. A major lesson learned was that developing a detailed, well-elaborated evaluation ToR with clear expectations for the evaluation, and assigning a qualified MCO staff member with strong RBM competencies and independence from project management contributed to enhancing the quality of the evaluation.

Natalia Galat, Programme Specialist
Multi Country Office for Central Asia
For example:

- Assess the **relevance** of UN Women contribution to the intervention at national levels and alignment with international agreements and conventions on gender equality and women’s empowerment.
- Assess **effectiveness and organizational efficiency** in progressing towards the achievement of gender equality and women’s empowerment results as defined in the intervention.
- Assess the **sustainability** of the intervention in achieving sustained gender equality and women’s empowerment.
- Determine the **impact** of the intervention with respect to gender equality and women’s empowerment.
- Analyse how **human rights approach and gender equality principles** are integrated in implementation.
- **Identify and validate lessons learned, good practices and examples and innovations** of efforts that support gender equality and human rights in area of work.
- Provide **actionable recommendations** with respect to the UN Women intervention.

**Defining evaluation criteria**

Evaluation criteria are the standard against which evaluation judgments are made. UN Women evaluation criteria are aligned with those of [UNEG](#), including criteria based on gender equality and international human rights principles. The standard criteria include: **relevance**, **effectiveness**, **efficiency**, **sustainability** and **impact** (where feasible). In addition, coherence, connectedness, coverage, coordination, innovation, and **gender equality** should be assessed, including issues such as **equality and non-discrimination**, **participation**, **inclusion**, **empowerment**, **accountability** and **social transformation**. A stand-alone criterion on gender equality and human rights is recommended to ensure a detailed analysis of these aspects. The selection of evaluation criteria for evaluations is dependent on the specifics of each evaluation and should be informed by the characteristics and context of the intervention in question. In case one or more of the standard evaluation criteria are not used, a rationale for the decision should be given in both the ToR and evaluation report. It is important to prioritize these criteria according to the information needs identified with stakeholders and the evaluable of those aspects.
Defining key evaluation questions

Evaluation questions contribute to further defining the objectives by relating to the purpose and criteria for the evaluation. The key questions outlined in the ToR will serve as the basis for the development of more detailed questions by the evaluation team. The questions should be precisely stated to guide the evaluator in designing the evaluation and in collecting information and data. Thus, the evaluability of the questions should be assessed according to whether or not the evaluation team will have the time, resources and information available in order to provide credible answers to the questions.

The evaluation manager will need to include evaluation questions that relate not only to the gender equality and human rights outcomes and impacts of the programme, but also to the gender equality and human rights dimensions of the planning, monitoring and implementation stages of the intervention. It is important to assess the effects of all interventions on gender equality and human rights, no matter the nature, focus or original intentions of the intervention. UN Women should also assess the inter-linkages between normative support work and its operational implementation and the progress of UN Women in supporting its UN system coordination role.

The evaluation manager should gain consensus on the evaluation questions, including the gender equality and human rights aspects, with stakeholders through the reference and management groups. The questions could be organized around the chosen evaluation criteria and the evaluation should present findings accordingly. Generally, three to five key questions related to each of the selected criteria will provide for a more focused evaluation.

For example:

• To what extent is the intervention relevant to the needs and priorities as defined by beneficiaries?

• To what extent is the intervention aligned with relevant normative frameworks for gender equality and women’s empowerment?

• What is UN Women’s comparative advantage in this area of work compared with other UN entities and key partners?

• To what extent were the expected outcomes achieved and how did UN Women contribute towards these?

• To what extent was gender equality and women’s empowerment advanced as a result of the intervention?
• What were the unintended effects, if any, of the intervention?
• To what extent was capacity developed in order to ensure sustainability of efforts and benefits?
• How will the benefits of the intervention be secured for rights holders (i.e. what accountability and oversights systems were established)?

**Defining the scope**

The scope of the evaluation describes what will be included and what will be excluded from the evaluation. Defining the scope provides an opportunity to weigh what is important to obtain from the evaluation against what is actually feasible.

The scope of an evaluation defines:

- **Timing**: When in the life of the programme is the evaluation being conducted (mid-term, end of programme, etc.)
- **Time frame**: Specific time frame in the life of the programme (the entire programme life or several strategic note periods, etc.)
- **Geography**: Whether or not it will cover the entire region where the programme has operated/provided services or selected areas
- **Programmatic or project coverage**: Whether or not it will include all aspects of the programme or focus on specific elements
- **Thematic coverage**: Whether or not it will include all aspects of a theme (e.g., ending violence against women, political participation, etc.), or focus on a specific sub-theme (e.g., domestic violence, gender advocates, etc.)
- **Limitations**: Limitations of the evaluation given the scope

The scope should take into account other existing or planned evaluations of the same subject to limit duplication and make efficient use of scarce resources. The relationship between the planned evaluation and other related evaluations should be described, including how information from these other evaluations may be used to guide the evaluator in designing the evaluation and in collecting information and data.

**Selecting the appropriate evaluation design**

One of the most essential characteristics of a gender-responsive evaluation is how it is conducted, that is, the process and methods that embrace gender equality and human rights principles. The evaluation design section of the ToR should clearly describe the chosen design and include the process and methodology. While the type of evaluation and, if possible, its design should be
considered at the planning stage, the ToR should specify details of the design. The regional evaluation specialist can be consulted when determining the most appropriate evaluation design and methods.

The design selected will frame the conduct of the evaluation and determine which methods are most appropriate. The evaluation manager does not need to be an expert on gender-responsive evaluation design or methods, as the evaluation consultant or team hired will provide expertise in this area. The ToR should, however, define preliminary thoughts on a general approach to be taken in the conduct of the evaluation. The evaluation design will depend on the purpose, objectives and key questions of the evaluation and on the nature of information available to the evaluator(s), such as indicators, baseline information and specific targets. The evaluation can be formative (forward looking), summative (backward looking), ex-post (including impact evaluations) or real-time (typically used in a humanitarian setting). The design should promote inclusion and participation by employing gender equality and human rights responsive methodological approaches such as those with a focus on utilization, empowerment or feminist approaches.

Gender-responsive evaluation applies mixed-methods (quantitative and qualitative data collection methods and analytical approaches) to account for complexity of gender relations and to ensure participatory and inclusive processes that are culturally appropriate. Even for impact evaluation, as UNEG guidance explains, “The emerging consensus in literature on impact evaluation appears to be that most questions can best be answered by “mixed methods”. This might involve a mix of both quantitative and qualitative methods, or a mix of specific approaches within either of the two categories. Furthermore, approaches which “blend”

---

19 While there are many types of evaluation design, the evaluation literature refers to three primary types of evaluation designs: a) experimental: involves random assignment, control group and before/after measurements; b) quasi-experimental: involves comparison group and after measurements, and may or may not involve before measurements; and c) non-experimental: no comparison group and measures change only at the end of the intervention. Typically UN Women evaluations employ non-experimental or quasi-experimental design; an experimental design is rarely used and, if it is, it should be used as one component of the overall evaluation, as it should be combined with mixed-methods.


methods, such as quantifying some aspects of qualitative data are also increasingly seen as valuable.”

The evaluation methodology should enable achievement of the evaluation purpose, be aligned with the evaluation design, and address the evaluation criteria and answer the key questions through credible and gender-responsive techniques for data collection and analysis (see Box 11 and Tool 10. Tips for employing gender-responsive evaluation methods).

The methodology section of the ToR should:

- Outline a wide range of data sources (e.g., documents, field information, institutional information systems, financial records, social media data, beneficiaries, staff, funders, experts, government officials and community groups)

Box 11. Gender-responsive evaluation methods

- Use gender analysis frameworks (e.g., Harvard analytical framework, gender planning framework, social relations framework, women’s empowerment framework)
- Draw upon feminist theory and methodologies
- Are appropriate and relevant to both women and men
- Are participatory
- Ensure collection of disaggregated data
- Understand the constraints and challenges of informants
- Explore gender roles and power relations
- Are context and culturally sensitive
- Emphasize mixed methods (quantitative and qualitative)

• Describe data collection methods and analysis (e.g., appreciative inquiry\textsuperscript{24}, most significant change\textsuperscript{25}, case study, survey, interviews, focus groups, observation, site visit, etc.) that will address gender equality and human rights issues; the evaluator will elaborate on the final rationale for selection and their limitations (see Tool 13. Advantages and disadvantages of data collection methods).

• Identify participatory tools for consultation with stakeholder groups and suggest a plan for inclusion of women and individuals/groups who are vulnerable and/or discriminated against in the consultation process and a plan for translation, as necessary

• Specify that the evaluator should detail a plan on how protection of subjects and respect for confidentiality will be guaranteed

• Specify that the evaluator should develop a sampling frame (area and population represented, rationale for selection, mechanics of selection, limitations of the sample) and specify how it will address the diversity of stakeholders in the intervention

• Specify that the evaluator should take measures to ensure data quality, reliability\textsuperscript{26} and validity\textsuperscript{27} of data collection tools and methods and their responsiveness to gender equality and human rights; for example, the limitations of the sample (representativeness) should be stated clearly and the data should be triangulated (cross-checked against other sources) to help ensure robust results

The evaluation process should outline the different phases of the evaluation, specify the key tasks evaluator(s) are responsible for carrying out, and include a schedule for completion. The details of the evaluation process depend on the selected approach. However, the following stages of the evaluation should be specified in the ToR:

• **Preparation**: This includes the stakeholder analysis and establishment of the management and reference groups, development of the ToR, and recruitment of the evaluation team

• **Conduct**: Inception report, stakeholder workshop, data collection and analysis

• **Reporting**: Presentation of preliminary findings, draft and final reports

• **Use and follow up**: Management response, dissemination of the evaluation products, and follow up to the implementation of the management response


\textsuperscript{25} Davies R, Dart J, ‘The most significant change (MSC) technique: A guide to its use’, United Kingdom and Australia, April 2005, available online at http://www.mande.co.uk/docs/MSCGuide.pdf.

\textsuperscript{26} Reliability is consistency in results using the same method (i.e., if the same survey is instituted several times it should give you similar results each time).

\textsuperscript{27} Validity refers to the degree to which the evaluation and its data collection tools are measuring the concepts intended to be measured; in other words, whether or not the tools are collecting the information they are intended to collect or measuring the right construct.
Ethical considerations for evaluation methods

One of the most important steps in assessing the different methods to be used by an evaluation is considering the ethical safeguards that can be employed to protect the confidentiality, dignity, rights and welfare of human subjects.

Through the evaluation, sensitive information may be collected regarding the violations of rights, which will require those affected to discuss these issues with the evaluators. UN Women and the hired evaluator(s) must assess whether or not certain methods may pose more harm than benefit to the affected individual or community. In order to safeguard those who participate in the evaluation, a clear plan for how information will be used, reported on and who will benefit from this information needs to be spelled out in the ToR. The plan should elaborate how informed consent will be obtained and make explicit that the names of individuals consulted during evaluation data collection will not be made public. In particular, a protocol for engaging those affected by violence should be elaborated in order to ensure the protection of rights and avoidance of further harm (see Box 12).

The UNEG Ethical Guidelines should be applied to the selection of methods for the evaluation and throughout the evaluation process. Following these principles is essential to ensure the inclusion of women, individuals and groups who are marginalized and/or discriminated against. In particular a gender-responsive evaluation must adhere to the obligations to participants:

- Respect for dignity and diversity
- Right to self-determination
- Fair representation
- Compliance with codes for vulnerable groups (e.g., ethics of research involving young children or vulnerable groups)
- Redress
- Confidentiality
- Avoidance of harm
Box 12. Evaluating violence against women and girls

Special precaution should be taken when the topic of the evaluation is on violence against women or includes vulnerable populations (e.g., survivors of violence, trafficked persons, etc.). UN Women must ensure that the rights of the individual are protected and participation in the evaluation does not result in further violation of their rights.

UN Women evaluations must ensure an explicit protocol is elaborated and adhered to for engaging vulnerable populations based on the following resources:

- World Health Organization (WHO), “Putting women first: Ethical and safety recommendations for research on domestic violence against women”, 2001
- WHO, “Researching violence against women: a practical guide for researchers and activists”, 2005
- WHO, “Ethical and safety recommendations for interviewing trafficked women”, 2003
- UN Women, “Virtual knowledge centre to end violence against women and girls”
- UNEG guidance document, “Integrating human rights and gender equality in evaluations”, Chapter 3
D. SELECTING AN EVALUATION TEAM

The ToR is the basis for soliciting and selecting evaluators. It defines the needs of the commissioning entity and specifies requirements. It should outline the skills, experience, qualifications and other relevant competencies that will be needed to conduct the evaluation effectively—whether by a consulting firm or by a consultant(s) (see Box 13). It has to specify the size of the team required and provide the estimated number of days required to undertake the evaluation. If resources allow, an evaluation team (i.e., more than one evaluation consultant) is ideal because they can provide a broader range of expertise and perspectives. For example, the team should hire one evaluation expert and one thematic expert and, to the extent possible, both should have gender analysis expertise. Large evaluation teams should be multicultural with gender balance and geographic representation. Evaluators are required to submit two or three examples of evaluation reports recently completed when responding to the ToR.

Box 13. Example of evaluator skills and competencies

- Experience in conducting gender-responsive evaluation
- Extensive knowledge of, and experience in applying, qualitative and quantitative evaluation methods
- Experience in gender analysis and human-rights based approaches
- A strong record in designing and leading evaluations
- Data analysis skills
- Excellent ability to communicate with stakeholders
- Technical competence in the sector or issue to be evaluated
- Process management skills, such as facilitation and communication skills
- Knowledge of the role of the United Nations and its programming at the regional and country level
- Language proficiency
- Country or regional experience
The recruitment and selection of external evaluation professionals will be governed by UN Women’s contracting policies and procedures. The following tips will help to ensure a strong evaluator(s) is identified:

- **Consider carefully all of the expertise** that may be required to conduct the evaluation based on the nature of the programme and its context. This expertise may not be found in a single individual but rather as the collective expertise of the evaluation team.

- **Consult key stakeholders** in the definition of skills required, criteria for selection, and consult the management group, including the regional evaluation specialist, on the final selection.

- **Ensure the ToR** clearly identifies requirements for evaluator(s) and indicates the expected time frame and deliverables. The ToR should contain enough specifications for a wide range of applications but should be cautious as to not over-specify so that the evaluator(s) hired have enough flexibility in designing an evaluation that best meets the needs defined in the ToR.

- **Undertake an open and competitive process** to recruit the evaluation team. The process should be impartial, fair and transparent, and there needs to be sufficient time allowed for the recruitment.

- **Engage local professionals**, as it provides better understanding of the local context and can be a catalyst for “buy-in” of the evaluation. It can also lend credibility to the evaluation process and recommendations. It is also an important means for strengthening national capacities in gender-responsive evaluation. Finally, local professionals will usually reduce evaluation costs because of travel costs. However, concerns regarding impartiality need to be carefully considered in their selection.

- **If a team is hired, consider the overall team suitability**. Will the individuals function well as a team? It is also important to ensure that the different skills and perspectives are balanced (see Box 13).

Locating the right expertise or combination of expertise is not easy and requires time and effort. Applicants should be given three to four weeks to prepare their proposals to maximize the number of potential applications from quality evaluators. Request support from the regional evaluation specialist to circulate the opportunity amongst evaluation networks.

Search the UN Women gender and evaluation consultant database.
Avoiding conflict of interest

Impartiality is an important feature of evaluation because it ensures credibility of the evaluation. In order to avoid a conflict of interest, those who are involved in the evaluation process must be impartial (i.e., they do not have a vested interest in the programme). For this purpose, as indicated at the beginning of this chapter, the programme officer responsible for the programme to be evaluated should not have individual decision-making responsibility in the evaluation process, including the ToR, final report and selection of the external evaluation team. Likewise, UN Women evaluations should engage an external evaluator who was not involved in the design or implementation of the programme.

Ethical conduct of evaluators

UN Women has developed a UN Women Evaluation Consultants Agreement Form for evaluators that must be signed as part of the contracting process, which is based on the UNEG Ethical Guidelines and Code of Conduct. These documents should be annexed to the contract. The UNEG guidelines note the importance of ethical conduct for the following reasons:

• Responsible use of power: All those engaged in evaluation processes are responsible for upholding the proper conduct of the evaluation.

• Ensuring credibility: With a fair, impartial and complete assessment, stakeholders are more likely to have faith in the results of an evaluation and to take note of the recommendations.

• Responsible use of resources: Ethical conduct in evaluation increases the chances of acceptance by the parties to the evaluation and therefore the likelihood that the investment in the evaluation will result in improved outcomes.

E. QUALITY ASSURANCE OF THE TOR

After drafting the ToR the evaluation manager should share it with the regional evaluation specialist for feedback. Then the ToR should be shared with the evaluation management group. The head of the respective office (country representative or regional director) should have the final approval of the ToR (see Figure 6 and Tool 1. Evaluation process standards for decentralized evaluation).

Sharing the ToR with the reference group will make good use of their collective knowledge and facilitate an agreed-upon understanding of the purpose of the
evaluation (key users and uses) and key elements of the evaluation, thus facilitating ownership. It will also serve to manage the expectations of key stakeholders and bring clarity to their role in the evaluation. In particular, UN Women should ensure the impartiality of the evaluation process by establishing distinctive roles between the evaluators, the commissioning party (UN Women) and the reference group in the ToR. The final ToR should reflect the shared understanding of the evaluation scope, methods and process among stakeholders.

The ToR should also provide the GERAAS report quality standards as an annex so that the evaluator is informed about the requirements prior to payment for UN Women reports (Tool 14. GERAAS evaluation report quality assessment checklist).

Figure 6. UN Women evaluation process: terms of reference
## TERMS OF REFERENCE (TOR):

Did the office appoint an evaluation manager (either the M&E officer or another staff member that is not involved in the programme management)?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Was the draft ToR shared with the regional evaluation specialist for quality review?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Was the draft ToR shared with the evaluation reference and management groups?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Was the final ToR approved by the country representative or deputy representative?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Did the M&E officer or focal point upload the final ToR to the GATE website?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

## Selection of Consultants:

Did the M&E officer/evaluation manager consult the regional evaluation specialist on the selection of the consultant/firm for the evaluation?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Was the final selection of the consultant/firm approved by the country representative or deputy representative?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
UN Women GATE: http://gate.unwomen.org

UN Women gender and evaluation consultant database: http://evaluationroster.unwomen.org

UN Women virtual knowledge centre to end violence against women and girls: http://www.endvawnow.org/

- Tool 1. Evaluation process standards for decentralized evaluation
- Tool 5. How to conduct an evaluability assessment
- Tool 6. Evaluation terms of reference template
- Tool 7. Evaluation product comment template
- Tool 9. Stakeholder analysis template
- Tool 10. Tips for employing gender-responsive evaluation methods
- Tool 11. Management group terms of reference template
- Tool 12. Reference group terms of reference template
- Tool 13. Advantages and disadvantages of data collection methods
- Tool 14. GERAAS evaluation report quality assessment checklist


- UNEG: http://unevaluation.org/
  - Ethical Guidelines
  - Code of Conduct
  - Guidance on Impact Evaluation
  - Integrating human rights and gender equality in evaluations
  - Norms for Evaluation in the UN System
  - Standards for Evaluation in the UN System
- UN Women: http://www.unwomen.org/en/digital-library
  - Evaluation Policy
  - Guide for the evaluation of programmes and projects with a gender, human rights and intercultural perspective
  - United Nations System-wide Action Plan
- UN Women intranet: https://unwomen.sharepoint.com/
  - Contracting Policies and Procedures
  - Evaluation Chapter of the POM
  - Evaluation Consultants Agreement Form
- Women’s empowerment framework: http://awidme.pbworks.com
  - Gender planning framework
  - Harvard analytical framework
  - Social relations framework
- World Health Organization (WHO): http://www.who.int/gender/documents/women_and_girls/
  - Ethical and safety recommendations for interviewing trafficked women, 2003
  - Ethical and safety recommendations for researching, documenting and monitoring sexual violence in emergencies, 2007
  - Putting women first: Ethical and safety recommendations for research on domestic violence against women, 2001
  - Researching violence against women: a practical guide for researchers and activists, 2005