Chapter 7. Conduct/Implementation of an HR & GE Responsive Evaluation

186. The purpose of this chapter is to underline the key elements that should inform any methodological choice in order to (1) understand if an intervention has been guided by and has achieved HR & GE principles and purposes, and (2) ensure that the evaluation process itself adopts HR & GE principles and purposes.

187. While, as mentioned above, it belongs to the evaluation commissioner to broadly define how the evaluation should be conducted, it is one of the first tasks of the evaluators to define “what information is required to answer [the identified] questions, from whom and how the information can best be obtained. [Decisions also need to be made on] how the information collected should be analysed and used.”

188. It is also at this stage that the HR & GE issues will have to be identified in line with the general guidance contained in this document, challenges outlined and the appropriate methodology defined. This part of the evaluation process should be informed by the evaluability study, where opportunities and challenges regarding HR & GE elements in the evaluation are assessed, and by the stakeholder analysis. The outcome of all these reflections will generally be outlined in an inception report or similar document.

189. To ensure the credibility and usefulness of the evaluation, the evaluation manager and the evaluators “must ensure that fieldwork meets evaluation method standards for gathering evidence to support findings and recommendations on the intervention’s contribution.” This entails that existing strategies and methods should be tailored to respond to specific HR & GE questions. When deciding among different methods and instruments, it is useful to question, in particular if the selected method(s) or tool(s) will:

- Adequately answer HR & GE issues by detecting meaningful changes and the contribution of the intervention to them in terms of enjoyment of rights, empowerment of rights holders and capacity of duty bearers;
- Be suitable for the populations and individuals that will be involved (in particular, if cultural and security issues are taken into account); and
- Be appropriate to involve all the key stakeholders, without discriminating against some groups or individuals, and allow for guaranteeing the meaningful participation of all stakeholders.

---


190. It is important to note that *the designed methodology needs to be flexible, taking into account that it has to be adapted to the intervention and country contexts*. This is particularly true of HR & GE responsive evaluations, since they might tackle sensitive issues and be carried out in highly politicized or insecure contexts. Evaluators need to adapt their methods to the risks of the persons involved – directly or indirectly – in the intervention and/or in the assessment process, as well as their own risks. These risks could be political, social or security in nature: think about the threat for a regime opponent to be seen talking with foreigners or the danger that in certain circumstances a woman may face in traveling to meet with the evaluators. A good knowledge of the social, historical and political context and constraints is needed.

191. In order to mainstream HR & GE in the evaluation process and to capture relevant HR & GE intervention results, the necessary amounts of funds, time and human capacity should be allocated. The evaluation budget, though, is usually decided at the design stage of an intervention. The evaluators’ review suggested above might then lead to renegotiations of budget allocations and/or additional resources to ensure HR & GE data generation and in-depth analysis. In this process, though, it is important to take into account that additional evaluation resources might not always be the answer, especially for interventions with a low evaluability of the HR & GE dimensions.

### 7.1. Data collection

192. Whenever possible, data should come from more than one category of respondents and more than one source. For example, if duty bearers report increased success in responding to rights holders’ claims and in protecting rights, this may be confirmed through records of decisions, or asking rights holders if they have noticed any changes in the negotiation processes with duty bearers and in their enjoyment of rights. If statistics report an increase in women’s income, the evaluation should ask women and their families whether they have observed this increase in their daily lives and how they have used the income. Local businesses can also be asked whether they have perceived an increase in purchases by women, and local banks can be asked whether they have noticed an increase in savings made by women. Triangulation completes and enriches findings.

193. *A combination of data collection methods is usually recommended to gather and analyse information, in order to offer diverse perspectives to the evaluation, and to promote participation of different groups of stakeholders.* Using a mixed-method approach usually helps improve the evaluation quality overall but has also emerged as being effective in capturing and integrating HR & GE perspectives and principles into evaluation processes, in particular transparency, non-discrimination, participation and inclusion. It provides the opportunity to carry out exercises to ensure that the voices of women, those most likely to have their rights violated, or those marginalized and/or discriminated against are heard and taken into account during the evaluation.

194. Using mixed methods also serves to validate the findings obtained from diverse methods through iterative testing and parallel, sequential or multilevel analysis. This is an effective mechanism to build defendable conclusions, which is of particular interests in evaluations concerning sensitive and sometimes questioned issues.
Box 20. Good Practice in Design: Using a Mixed Method Approach

This study examines the different social capital profiles of women and men in Australia, testing the hypothesis that women’s caring and community-based responsibilities may constrain their civic and political aspirations. It sought to explore social capital in two different ways: first, to map the different patterns of participation based on gender, and second, to explore how the role of “mother” alters both the activities women become involved in and the reasons for this.

While this example is a research study rather than an evaluation, it was chosen as a good practice case because it consciously set out to use a mixed method approach, and illustrates that this approach is feasible given conceptual clarity and adequate capacity, time and resources. The study locates itself within a transformative research paradigm, which is seen as providing a framework for addressing issues of social justice in the research process.

The transformative paradigm recognizes that voices of those who are disenfranchised on the basis of gender, race/ethnicity, disability or other characteristic can be excluded in research. Within this paradigm, mixed methods are preferred to highlight issues of need (quantitative data) and to give voice to these issues (qualitative data). Feminist research that draws on evidence from a variety of sources is more likely to be seen as valid and reliable and is thus more likely to be heard in the policy arena – and the same case could be made for feminist evaluation.

The study’s author recognizes that in large quantitative research, women’s voices as an oppressed group have remained unheard, while with qualitative research, problems with poor representation and a tendency to overgeneralize need to be highlighted. The researcher used sequential mixed methods sampling in two stages. In stage one a large sample was chosen through simple random sampling, with a questionnaire on social capital going to 4,000 people, and eliciting 1,431 responses.

Participants who were interested in being interviewed for the second stage signed an agreement form sent with the initial questionnaire, and 12 respondents were then chosen for intensive interviews by cluster random sampling technique (where already formed groups of individuals within the population are selected as sampling unit). Quantitative data was analysed using standard statistical techniques, including multivariate analysis. Qualitative data was analysed using a model of narrative analysis, looking for plot, characters, metaphors, interpretations and cultural norms; how the stories compared and contrasted; and how the researcher was viewed by the participant. Findings from the quantitative and qualitative elements of the study were compared.

The author concluded that: “Despite a considerable body of literature devoted to social constructions of gender roles, there is little discussion in the social capital literature on the effect of gender. The power of a mixed methods research approach has been to build a comprehensive picture that challenges this lack of attention in the social capital literature.”

195. The extent to which an evaluation will be able to combine methods to evaluate HR & GE processes and results partly depends on resources, time and expertise. But for virtually any evaluation, it should be possible to include at least some elements of a mixed-methods approach. Box 20 brings a practical example of a multi-donor/multi-method evaluation process addressing HR & GE.

**Box 21. Reaching Women and Individuals/Groups Who Are Marginalized and/or Discriminated Against in a Country Programme Evaluation: Using the ‘Snowball’ Technique**

The ‘snowball’ technique, or respondent-driven sampling, where existing study subjects recruit future subjects from among their acquaintances, is one means of identifying women and individuals/groups who are marginalized and/or discriminated against when developing a research sample. Although subject to possible biases (e.g. initial respondents may select friends or relatives as the future subjects or their selection may represent their own biases based on class, race, ethnicity, caste, gender, etc.), it is a rapid and cost-effective means of identifying usually invisible groups.

This technique was used in an evaluation of the WFP India Country Programme (2007). In its planning documents, WFP strongly emphasized that its target was to reach some of the country’s poorest districts, and within them the most food insecure households as the primary target group, in particular women, girls and infants.

During the evaluation, a form of ‘snowball’ methodology was used at the village level to reach women and individuals/groups who are marginalized and/or discriminated against. A focus group discussion with up to 20 individuals was held at the start of the visit to each village, to have an overview of the core programmes. During that discussion, three to five of the poorest households in the village were identified. These households were then visited to assess the extent to which WFP support had reached the most food-insecure households. The evaluation team also ensured that approximately 50 percent of respondents were women. This methodology was useful in determining how effectively WFP had been able to reach its core target group.


196. Within a mixed method approach, each data collection method or tool can then be adapted to integrate HR & GE dimensions.

197. When using *samples* (such as purposeful sampling, theoretical sampling or snowball sampling), the selection of the sample is crucial since it can affect the credibility and technical adequacy of the information gathered. For HR & GE responsive evaluations, it is important to ensure the representativeness of stakeholders transparently and without discrimination (see for example in Box 21 the application of ‘snowball’ technique to reach women and individuals/groups who are marginalized and/or discriminated against in a country programme evaluation). Evaluators should also consider that comparisons between large groups, be they ethnic, sexual, socio-economic or geographical groups, could hide considerable diversities within the group.
198. If the amount of information reviewed or data collected is too limited, the findings may be questioned. If budget concerns or time constraints limit the number of respondents, or if the number in some categories is very small (for example, only a few representatives of one affected ethnic minority can speak with evaluators), the findings need to be validated by a larger group, or through triangulation. The sampling strategy also needs to address the inclusion of women and men in diverse stakeholder groups. In dealing with such diverse samples, the data collection strategy may need to contemplate several collection methods and alternatives to reach those women and individuals/groups most marginalized and/or discriminated against. Sometimes, even representative samples are too small to capture diversity within the total population; it will then be prudent not to generalize findings and not to report in terms of percentages.

199. HR & GE data disaggregation should be favoured. While sex disaggregation is the most common form of disaggregation across the UN, a HR & GE responsive evaluation should go beyond that. Understanding the nuances within groups as well as any form of exclusion (such as age, disability status, ethnic origin, place of residence, sexual orientation, social class or income group, etc.) will offer the evaluators a much broader view of how the intervention affects all the stakeholders involved. Data disaggregation can be a powerful ally to triangulation, as the diversity in responses obtained can prepare the ground for cross-examination, using other methods and by asking different sources. Note needs to be taken that extensive disaggregation of the data, especially if broken into multiple smaller subgroups, could be questionable in regard to generalizability.

200. Evaluators can make good use of existing national or international data sets (on employment, income, vulnerability, disease, mortality, human rights violations, etc.) to compare and confirm or refute findings. The use of these data, nevertheless, should be undertaken with an understanding of their possible limitations and constraints in representing local reality. It may be useful and efficient to test findings with a diverse panel of experts, who can corroborate or suggest other interpretations. This may be particularly useful for impact evaluations but also for small evaluations where fieldwork is limited.129

201. Existing national and international data sets: Evaluators can tap into a wide range of secondary data sources to better understand the HR & GE situation in the country, region or community they are researching, and to support their conclusions through triangulation. Data generated by governments, international organizations, academia and civil society can be found in a myriad of analyses and documents, including:

- Data produced by national and international statistics institutes. These data can concern population statistics, the implementation of international human rights obligations, violence, socio-economic indicators, or the situation of women and individuals/groups who are marginalized and/or discriminated against. Some of them might contain disaggregated data, according to considerations such as sex, age, ethnic communities, etc. Many countries have started working on the adoption of indicators, including indicators on compliance

129 On the human rights indicators project, see <ohchr.org/EN/Issues/Indicators/Pages/HRIndicatorsIndex.aspx>; USAID is financing the Demographic and Health Surveys Programme to collect representative data on population, health, HIV and nutrition in over 90 countries. See <www.measuredhs.com/>. 
with international human rights and gender equality commitments, to be monitored by
national institutes for the benefit of all branches of the Government. National surveys may
provide useful quantitative data regarding demography (mortality and morbidity rates),
employment, income, violence, health, sexual and reproductive rights, etc. Evaluators may
also benefit from qualitative research inquiring into cultural mentalities and behavioural
attitudes related to women, gender relations and individuals/groups who are marginalized
and/or discriminated against.

- **Data produced by governments to respond to international treaty-based or Charter-
  based human rights bodies.** Periodic reports submitted by States Parties to international
treaties-based human rights bodies and the concluding observations/recommendations of
these committees contain summaries of shortfalls vis-à-vis the implementation of interna-
tional human rights obligations, as well as capacity gaps in implementing HR. For exam-
ple, national CEDAW reports contain important analyses on the situation and progress of
women’s rights. Charter-based bodies, such as the Human Rights Council and its Special
Procedures, also offer a wealth of information. In addition, the Universal Periodic Review
of the Human Rights Council may provide useful contextual information to the incidence
of human rights abuses. Special HR Rapporteurs, representatives and working groups also
issue reports that can be extremely useful for evaluations.

- **Data produced by international organizations.** Situation analyses such as the CCA, pre-
liminary analyses for poverty reduction strategies, and HR & GE analyses carried out as
background studies to design new interventions are an important source of information
that should be integrated into the background document analysis for an evaluation. HR
monitoring reports (for example, as carried out by UN organizations such as OHCHR) and
specific UN agency reports, such as UNDP’s Human Development Report, may also offer
important data. Bilateral cooperation agencies may also commission research on the HR &
GE situation, which should be considered as well. National reports on the MDGs will also
provide specific information on the situation of women and children, and on other groups
likely to have their rights violated.

- **Independent reports and research studies produced by academia and national and inter-
national civil society organizations.** Evaluators should look at the existing body of quan-
titative and qualitative research on HR & GE, such as studies commissioned by academia
and civil society. They may provide alternative points of view and inquiry areas that can
complement the information obtained in the evaluation. Apart from research, CSOs often
collect, systematize and make available information on human rights violations, for exam-
ple, in the form of databases.

- **Nationally and locally produced reports in the context of the intervention.** Programme
reports and other documents produced by partners and stakeholder organizations addressing
HR & GE issues and indicators can offer invaluable insights into the situation of the
particular communities and groups affected by the intervention. Programme monitoring
reports are an essential input to evaluations.
There is a very wide range of other data collection methods and tools currently in use in evaluation practice. Some are particularly geared towards embracing HR & GE principles and are therefore examined here in more detail.

Document review and analysis

In order to integrate HR & GE issues into a background document analysis, the evaluators should first look for specific information on HR & GE in the intervention being evaluated, such as:

i) evidence of a HR & GE analysis at the design stage (including HR & GE indicators);
ii) evidence of a detailed and inclusive stakeholder analysis, including women and individuals/groups who are marginalized and/or discriminated against;
iii) evidence of quality engagement and participation of stakeholders in the various steps of implementation;
iv) information on various stakeholder groups collected during monitoring activities;
v) evidence of how HR & GE were addressed by the intervention, and the results achieved in the area. Additional documents could also be useful, such as:

i) organizational policies, system-wide policies and mandates, agreements, etc. on HR & GE;
ii) literature produced by programme partners and other organizations that may inform the assessment of HR & GE in the intervention.

Focus groups

Focus groups are highly relevant for HR & GE responsive evaluation as they can encourage women and individuals/groups who are marginalized and/or discriminated against to express their views more openly than through conventional survey methods. However, they do not automatically guarantee that participants will use the opportunity unless they are carefully designed and facilitated with this in mind.

Tailoring them to address HR & GE issues involves:

- **Paying special attention to the constitution of groups.** The choice of how to constitute a focus group needs to rely on the evaluation questions and stakeholder analysis defined in the beginning of the evaluation process, but also on factors such as the context of the intervention, the practical feasibility to disaggregate participants and common sense by the evaluator on what would be a better mix in each particular intervention.

- The constitution of the groups will have a significant influence on the extent to which participants feel safe to participate and communicate their ideas. This is highly relevant to consider when dealing with HR & GE issues – participants can be seriously affected for having made statements at the wrong time and in the wrong place. Evaluators should be extremely conscious of what the risks can be, particularly in certain countries and situations.

- The evaluator has the option to seek disaggregation by sex, age, social position, income, sexual orientation, category (rights holders/duty bearers), disability, etc., in order to investigate in-depth the meanings attached to a given phenomenon by a subgroup of population. Alternatively, creating mixed groups with careful facilitation may also provide important insights into group dynamics, and how different groups relate to each other. A mixed focus
group may also have the advantage to garner different perspectives and generate discussions that may not arise in a very uniform group. It can also help examine whether consensus exists among different parties. Nevertheless, consideration should always be given to the possible dangers of bringing together individuals in unjust relations of power (e.g. duty bearers and right-holders; ethnic majorities vs. ethnic minorities, etc.)

- **Facilitating responsively.** Before starting the focus group, it is important to seek information to help understand the context, the relationships between individuals and groups, the power dynamics, and how HR & GE issues affect the different individuals and groups represented in the focus group. This knowledge should help to guarantee an adequate group interaction during facilitation, and later to inform the analysis of the focus group discussion.

- **Carefully considering language and culture issues**, as many stakeholders may not be fluent in the main language of the evaluation, or may have different understandings of concepts discussed. In this case, field testing of the interview questionnaire/guide or advance cognitive interviews with individuals from various language/cultural groups could be helpful. The support from a national consultant might also, in certain circumstances and conditions, be recommended. National consultants should be used insofar as possible, but language, ethnic group or culture and sex of the interviewers must be carefully matched to the characteristics of the participants in the focus group, to avoid conflicts and barriers to communication. For example, in many contexts, a man may not facilitate a focus group of women; the inverse situation might also be true in other circumstances. Furthermore, when discussing HR issues, it is necessary to consider that national consultants might pose some problems, for example, if they belong to a specific ethnic group or to a certain class, or have specific family or institutional associations. The problem can be for the interviewees but also for the national consultants themselves (for example, travelling to certain areas for interviews might be very dangerous for them).

- **Promoting progress on HR & GE.** While focus group discussions’ primary purpose is to collect data for evaluating a specific intervention, it also provides a space for stakeholders to have a dialogue, exchange views and gain a better understanding of different perspectives and ways in which an intervention can have a diverse effects (positive and negative) on different stakeholders, which is linked to the larger social, economic and cultural context and gender relations. In this way, focus group discussions can contribute to attitude changes that are key to addressing inequality and discrimination.

**Individual interviews**

206. Often, women and individuals/groups who are marginalized and/or discriminated against are not involved in the evaluation process. The following are the most common reasons:

- They may not be able to express themselves freely when consulted because of social pressure, e.g. from elites, the community or their relatives;
- They may be persons with disabilities (e.g. deaf or blind people, people with intellectual disabilities) whose accessibility to the evaluation activities and sites may be difficult;
They may be illiterate or less fluent than others in the language used in interviews;

- They may not be allowed to speak, use their own language or be represented in public meetings or community consultations;

- Women may have less time at their disposal because of their productive and reproductive tasks, or may defer participation to males in observance of existing gender norms.

207. In order to address HR & GE issues through interviews, the evaluator should:

- **Make sure that the sample selected for individual interviews adequately reflects the diversity of stakeholders of the intervention.** For advocacy, normative or broader policy work, other types of persons also need to be included. Special attention should be paid to the inclusion of women and individuals/groups who are marginalized and/or discriminated against, who may have been forgotten or left out of discussions and decision-making in the intervention. The inclusion of women should be also sought. The selection of interviewees should be closely related to the evaluation questions and the stakeholder analysis, but also rely on a good understanding of the context. A national consultant could provide very important insights during this stage (with all the reserves expressed regarding national consultants in other part of the present document).

- **Consider language and translation needs.** This could represent a difficult issue while tackling HR & GE issues. There is a need to adjust the questionnaire to respondents who are illiterate or have low education levels, and make sure that all are able to understand the questions. It is also necessary to avoid using technical terms that may not be clearly understood by the respondents and might have different meanings and understanding. Human rights and gender questions must be adapted to the political, social and idiomatic contexts. In cases where the evaluation is being carried out in a local language, resources should be provided for translation or interpretation. Translation could also be challenging, not only because of language gaps but also because translators might distort the content of what is being said (for political or other reasons) or intimidate the interviewee (if they belong to specific ethnic or national groups for example). Careful selection of the interpreters is therefore required.

- **Consider practical measures** such as timing the interviews to fit home obligations, choose physically accessible venues, provide financial support for interviewees travel costs related to participation, etc.

- **Make sure that safeguards are taken to ensure that interviewees will not be negatively affected** by providing their honest views on HR & GE issues. If this is not possible, then the interviews should not take place; not include certain questions; and/or the possible danger made clear to the interviewee for him/her to decide on participation.

- **Respect confidentiality.** Ask permission to quote their words. In some cases, words or sentences may identify the person, even if their name is not in the report. In these cases, be honest about the confidentiality challenge and only quote interviewees if they agree with it. Even so, evaluators should use their wording with caution since the interviewee might not be fully aware of the consequences their words might bring to them. Use common sense to assess the
context and determine what the risks could be for the interviewee. Attention must also be paid to the list of persons interviewed provided in the report. In some contexts, such a list should not be included or be limited to broader information (institutions, origins of the interviewees, category of the persons interviewed – for example “three patients in hospital B”).

- **Make sure to understand how each interviewee is affected by HR & GE issues**, for example by asking specific questions as to how they see gender relations in their community, how they are affected by the practice or behaviour of duty bearers and by rights violations, what changes they have seen in the HR & GE situation in their community and what these changes have meant to their lives in practice. In some contexts, these questions might be highly sensitive and need to be properly tailored (for examples of possible questions to ask in relation to gender equality results see Table 14).

- **Make sure to ask specific follow-up questions on HR & GE during the individual interviews**. For instance, if respondents are discussing issues such as the creation of local organizations, make sure to ask questions such as the effect of these initiatives on gender relations, and their implications for the enjoyment of rights.

**Case studies**

208. A method that can be adapted to support the integration of HR & GE dimensions in evaluation is the use of case studies. This is a widely used social science technique that may be particularly helpful for highlighting the experiences and concerns of women and other groups likely to have their rights violated, or to study the effect of a particular policy on rights holders, or to analyse the behaviour of duty bearers. Case studies are context-specific and can help enrich the evaluation by providing a detailed analysis of specific instances such as events, institutions, policies, or by telling a story that may elucidate a particular situation. They are also particularly useful to describe good practices in an intervention.

**Box 22. Case Studies: A Feminist Perspective**

Reinharz (1992: 167-8), writing from a feminist perspective, notes that case studies are written “to illustrate an idea, to explain the process of development over time, to show the limits of generalizations, to explore uncharted issues by starting with a limited case, and to pose provocative questions. For example, a carefully chosen case can illustrate that a generalization is invalid. For this reason studies of the exceptional case have great heuristic value. Although they cannot establish a generalization, they can invalidate one and suggest new research directions. The exceptional case is valuable for feminist action, as a positive model to emulate or as a negative model to avoid.”

209. An example of a good practice case study from a WFP evaluation in Southern Africa is given in Box 23. This example addresses HR & GE issues by posing questions related to the right to food of families who are marginalized and/or discriminated against, by putting a woman in the centre of the analysis, and by empowering this woman to understand and interpret her own situation vis-à-vis the food security issue in her community.

The following case study is taken from the 2002-2003 WFP Real-Time Evaluation report that covered the six countries included in the Southern Africa Regional Emergency Operation (Lesotho, Malawi, Mozambique, Swaziland, Zambia, Zimbabwe) and the Regional Bureau. The evaluation included: document review; interviews with programme staff; focus group discussions with the project participants at final distribution points in each country; and household visits where in-depth semi-structured interviews with the population targeted were undertaken. In addition, an ad hoc ‘sentinel site’ study was undertaken in Malawi and Zambia, where families were visited on subsequent missions for an update on progress and an insight into the impact of the operation. It is this last method that provided the case study material below.

The household of Ambu, an elderly widow who lived with two of her daughters, their children and an orphaned grandchild in a small village in Ntcheu district, Malawi, was visited three times during the food emergency. Three of Ambu’s children died in their 20s, probably of AIDS. Rose, a daughter in her early twenties, is often sick and weak. Her husband left her many years ago. Dorothy, the other daughter, has a husband and four children, but the husband provided little to household income.

During the first house visit most family members appeared weak. Although the surroundings of the clay house was swept clean, Ambu’s field was in a depressing state. She had grown some maize but the harvest was poor and she ate much of it while it was green. By June, her own production was eaten up. She managed to get food (normally for one meal a day) through begging from villagers, gifts from a third daughter who is married to a teacher, and from occasional daily work of Rose with a farmer in the village who grows tobacco.

Ambu was then selected as a beneficiary in the first round of food distribution. The implementing NGO Africare managed an orderly distribution, based on beneficiary lists prepared by the village committees in line with selection criteria giving priority to households with orphans, etc. Ambu received a 50-kg bag of maize. She should also have received pulses and corn-soya blend but donations were not available.

At the time of the second visit, six months later, all family members are there and look a bit healthier. Rose had found work in a nearby village for about one month and was paid in maize meal. Ambu received the UK Department for International Development-sponsored seed ‘starter package’. She planted maize but is now waiting for rain. The growth chart for the grand-daughter shows regular growth. Ambu’s household eats currently two meals of maize meal with leaves from the Baobab tree. Asked whether she received this month’s food aid ration, Ambu says yes. In reality she did not. She dropped off the list of beneficiaries. The village chief explains that he had to rotate ration cards, which he keeps. Cards for 22 eligible households out of a total of 130 households in the village are far too few to meet the needs of the poorest households.

At the time of the third visit, five months later, Ambu’s name is back on the register of eligible households. But she did not get the monthly food ration. Only once, in March, did she receive a bag with 25 kg of maize. Africare staff explain that Ambu’s ration has been regularly picked up by a boy that had been identified as her representative. Ambu’s family is complete although Rose has been sick and weak again. Rose’s six-year old daughter has dropped out of school because she lost her schoolbook.
Surveys

210. Surveys are the most common tool for collecting standardized information from a large number of people in an evaluation, in particular target and control groups.\textsuperscript{130} In addition to already discussed issues regarding interview procedures, the inclusion of HR & GE issues implies adapting some aspects of survey procedures. For example, it calls for the design of specific questions, for particular techniques to interview the selected persons, and for careful analysis of potential biases (for instance, to understand why interviewees refuse to answer or, to the contrary, are keen to respond). It also involves:

- \textit{Making sure that the survey includes specific HR & GE questions and enables disaggregation of the data collected.}

- \textit{Paying particular attention to the format and language of the survey.} It is important to consider alternatives to address HR & GE questions and interpretation issues (see above subsection on interviews).

- \textit{Creating different questionnaires for different stakeholder groups.} While it is important to ensure that at least some of the questions are comparable in content (to inform the subsequent data analysis), it could be key, in certain circumstances, to address the specific issues and interests of the various stakeholder groups through tailored questions. This option needs to be well analysed, since developing several questionnaires might come at high costs and generate statistical problems.

Field observation

211. Field observation is a very effective, and sometimes crucial, tool for gathering information on HR & GE. The observation of the intervention activities and ongoing dynamics, and direct interaction with people or groups involved, allows understanding of aspects that might not surface when applying other methods. It can complement information obtained from other sources. This is particularly relevant when HR & GE dimensions of an intervention are culturally or politically sensitive. Field observation is a productive tool to:

- Formulate questions that can be posed in subsequent interviews;
- Examine the project’s physical and social setting, staff and clientele characteristics, group dynamics, and formal and informal activities;
- Become aware of aspects of the project that may not be consciously recognized by participants or staff;
- Learn about topics that programme staff or participants are unwilling to discuss; and
- Observe how project activities change or evolve over time.\(^{131}\)

212. Field observation needs to be carefully prepared to achieve its purpose and to avoid violating cultural or social norms, especially when considering HR & GE issues. As mentioned above, risk factors also need to be carefully weighed.

Training and use of local stakeholders to act as evaluators

213. Another option for HR & GE responsive data collection in a field situation is the training and use of local stakeholders to act as evaluators and to obtain further information, especially in those cases where there is a large sample size or geographical area to take into account. While such an action may appear to be somewhat counter-intuitive, this has proved invaluable for evaluation processes, as well as in the generation of results.

214. While the use of this methodology depends upon resources for the evaluation as well as the competence of the ‘local evaluators’, it has positive advantages. In the following example (Box 24), as women with the same cultural and linguistic dynamics were conducting the interviews of the women who participated in the project, there was a built-in comfort level between them, which led to more in-depth elaboration on results than it may have been possible to obtain otherwise. This methodology also empowers the ‘local consultants’ to gain confidence in their abilities, and have a direct input into the evaluation process. The one drawback of this methodology is the prospect for bias on behalf of the ‘local consultants’ when interviewing their peers. In the following example, this was somewhat mitigated by ensuring that each of them went to a different geographical area than that of their own cooperative.

**Box 24. Use of Local Stakeholders as Evaluators**

An evaluation of women’s credit cooperatives was conducted in Nepal. The cooperatives, which served as vehicles for democratic awareness, were spread all over the country, making it impossible to visit even a small representative sample in the two weeks allotted to the field mission. The solution was to organize at the beginning of the field mission a round-table training session for representatives of the cooperatives selected from around the country and put forward by their own cooperatives. The evaluator formulated all evaluation questions beforehand, then held a session with the women to vet the questions as well as to train them in how to conduct interviews with cooperative members, take notes, and ask follow-up questions based on certain responses. Each ‘local evaluator’ then went into the field and carried out the interview process with one or two cooperatives that were not their own, based on a set template. At the conclusion of the field mission, the evaluator met with the ‘local evaluators’ to discuss their findings, which were then incorporated into the evaluation report.

**Table 14. Possible Questions to Ask in Relation to Gender Equality Results**

Oxfam (2002) sets out five dimensions in which change can potentially occur as gender equality is strengthened, which could be used as measures of results during an evaluation. For each of these dimensions, possible questions are suggested, which will support gender analysis and orientate the interview.

<table>
<thead>
<tr>
<th>Dimension and results level</th>
<th>Suggested questions</th>
</tr>
</thead>
</table>
| Have women and men achieved more equal participation in decision-making in public and private spheres? Process | • Has women’s negotiating power in economic decisions (e.g. use of resources, money, time) and other family decisions (e.g. number of children to bear, type of contraception, children’s education) been strengthened?  
• Do women enjoy greater participation in the political processes of their communities?  
• Has the influence of women on decision-making increased in relation to that of their male counterparts? |
| Have gender stereotypes and discriminatory attitudes towards women and girls been challenged and changed? Process | • Do men and women better understand how unequal power relations between them discriminate against women and keep them in poverty?  
• Is women’s unpaid and caring work better valued?  
• Have changes in the traditional gender division of labour occurred with men taking on more household and caring work?  
• Is greater value attached to girls’ education?  
• Is violence against women increasingly rejected by the public, especially by men?  
• Are more men taking action to challenge discrimination against women? |
Table 14. Possible Questions to Ask in Relation to Gender Equality Results (continued)

| Process/outcome | • Has women’s self-esteem and self-confidence to influence social processes increased?  
|                 | • Are women more able to exercise their capacity for leadership?  
|                 | • Are women increasingly organizing to strengthen their voice and influence?  
| Do women and men have more equal access to and control over economic and natural resources and basic social services? | • Has women’s control over natural and economic resources (land, household finances, equipment, other assets) increased?  
| Outcome/impact  | • Do women have greater access to paid work?  
|                 | • Do women achieve equal pay for equal work with men?  
|                 | • Do women share the workload more equally with men and have more time for themselves?  
|                 | • Do women and girls have access to health services on an equal basis with men and boys, and according to their gender-specific needs (e.g. reproductive health)?  
|                 | • Do girls enjoy equal access to schools with boys?  
|                 | • Has the school environment become safer for girls and the curriculum less gender stereotyped?  
| Do fewer women suffer gender-related violence? | • Has the intervention led to a decrease in violence against women?  
| Outcome/impact  | • Has the intervention caused or exacerbated violence against women, or the fear of violence?  
|                 | • Has the number of women suffering personal incidents or threats of violence in the community or household changed?  

7.2. Data analysis/interpretation

Throughout the implementation of the evaluation, there will be some degree of data analysis (e.g. during document review, interaction with stakeholders, consolidation of survey data, etc.). ‘Iterative’ testing and analysis is advisable, particularly human rights and gender analysis, as early analyses will show, for example, where data is missing, what the most interesting questions are, etc. It can therefore pave the way for further data collection that is more targeted. However, it is at the end of the data collection stage that evaluators have enough material to carry out a complete data analysis. Data analysis and interpretation involve technical issues that are outside the scope of this Guidance. The focus here is on key elements that will ease the way for incorporation of HR & GE perspective in this phase of the evaluation.
216. Ideally, the data analysis and interpretation processes should involve key stakeholders, including duty bearers, rights holders, and within these two groups women and individuals/groups that are marginalized and/or discriminated against.

217. Generally, in HR & GE responsive evaluations, analysing data entails several or all of the processes below (some of the steps are not HR & GE specific but apply and are relevant to HR & GE analysis):

- **Comparing the data obtained with existing information on the situation of HR & GE.** This step allows the evaluators to establish whether most of the data collected during the evaluation confirms or refutes trends and patterns already identified. It also allows evaluating what gaps have been filled, and what new information has emerged.

- **Processing data from surveys.** When processing survey data, evaluators of HR & GE responsive evaluations should identify trends, common responses and differences between groups of stakeholders (including duty bearers and rights holders), disaggregated in different ways, such as sex, age, place of residence, belonging to minorities, disabilities, gender identity, etc. When correctly administered, survey data can be analysed in terms of cause and effect in the context of a specific theory of change, e.g. sex can be an explanatory variable for levels of poverty or ethnicity for levels of participation.

- **Making sure that an adequate understanding of the context, relationships, power, etc. informs the analysis of data collected in interviews.**

- **Comparing data obtained from different sources.** At this stage, it will be possible for the evaluators to triangulate information, and check whether there are similarities and/or discrepancies in data obtained in different ways and from different stakeholders. This comparison can also help to understand how different stakeholders are positively or negatively affected by the intervention.

- **Comparing individual stories and case studies with general information.** This is when the evaluator identifies the context behind the numbers, and the exceptions to the rule – which, as we have seen, might be particularly meaningful in terms of HR & GE. Individual stories and case studies may confirm trends obtained from quantitative analysis, and may also provide examples of how these trends are reflected in people’s lives. Or they may demonstrate that, even if a particular trend emerges, it is not reflected in the same way to everyone.

- **Comparing the results obtained to the original plan.** This is part of any UN intervention that follows the principles of RBM. The findings of an evaluation need to be compared with the original plan for the intervention, including its intended results and indicators. The evaluators should also ask whether the results framework has been sufficiently updated over time to reflect changes in the context of the programme. For HR & GE responsive evaluations, working with disaggregated data at this level is key, as it will allow the evaluator to probe whether the results are the same for everyone, or whether they benefited some more than others.
A particular aspect of analysing data is raised by the analysis of policies and programme strategies, including HR & GE policies. Several UN and international cooperation agencies have developed assessment tools in order to register progress in these areas. UNICEF and FAO present interesting examples. UNDP/UNFPA can also be cited as they have developed gender markers.

The Rights and Results Assessment Tool, set out in Table 15, was developed for UNICEF’s evaluation of its gender policy and is a generic tool that can be used to assess changes in the enjoyment of rights through programme-level interventions. The tool presents a scaled rating system for each of its components, including how far the intervention has promoted gender equality and contributed to meeting different institutional mandates. As an example, under section 1A the evaluator rating the programme determines whether the results planned in the intervention were at the level of women’s strategic interests or practical needs. As another example, for column 3 on results achieved, the evaluator is asked to determine whether gender equality results were partly or fully achieved, or if gender equality results surpassed expectations and objectives.

---


133 Table 15 offers an illustration of only one rating scale related to one type of gender equality result; the rating scales for all eight areas of the Tool can be found in the UNICEF evaluation.


135 The rating instructions are the following: 0 = no change; i = Change at practical needs level, i.e. change in material well-being and basic needs (short term, immediate changes related to gender gaps in basic needs); ii = Change at strategic interests level, i.e. change in the structural causes of gender inequality, critical awareness, advocacy, increased capacity for rights, participation, etc. (long term, social and capacity change leading to transformation of gender equality situation.)
220. Another tool developed to evaluate gender mainstreaming in programming is the six-point assessment tool, implemented by FAO in its evaluation of Gender and Development (GAD). The tool uses the following descriptors to assess public goods developed by the agency:

- **Relevance for GAD**: extent to which the GAD approach (mainstreaming GAD) is a relevant issue in the understanding and management of the topic treated;

- **Technical quality of GAD contents**: extent to which GAD issues are mainstreamed through the whole document with contributions aware of state-of-the-art discussions;

- **Innovations on GAD**: extent to which the document makes an innovative contribution to understanding of GAD issues;

- **Potential impact as advocacy tool**: extent to which the document is written with well-chosen case studies, and awareness of target audience and potential controversial aspects;

- **Potential impact as capacity development tool**: extent to which a clear argument and well-chosen case studies are coupled with either capacity-building materials or directions towards such materials;
• **Links between GAD and social inclusion**: extent to which GAD issues are mainstreamed into discussions of social inclusion with concrete suggestions as to successful integration.

221. Once the data is analysed, the evaluator will need to interpret the findings, moving to more detailed questions on finding causal links and making inferences. Taking a HR & GE approach, data should be interpreted if possible through multiple lenses, including for example sex, socio-economic status, ethnicity and disability. Groups most likely to have their rights violated are often subject to multiple forms of discrimination, and it is important to understand how these different forms intersect to deny rights holders their rights. Cultural sensitivity is needed in data management as in all other elements of evaluation practice.

222. The level of interpretation depends on the evaluation focus and on the level of resources available. Here are some suggestions:

- For all types of **interventions where a high level of resources** is available for evaluation, data interpretation involves assessing how power relations, including gender relations, have changed as a result of the intervention, and how the intervention brought about structural changes in these relations and in other human rights issues. This implies understanding the underlying causes of the development challenges tackled by the intervention, and to what extent these causes have been addressed. A detailed human rights and gender analysis can be carried out. For example, discriminatory cultural practices may have stopped, ethnic minorities may be voted into political office, minimum wage levels may be introduced and enforced, or the right to food may be ensured for women and individuals/groups who are marginalized and/or discriminated against. The evaluator will need to look in detail at what factors have contributed to or hindered these changes. Evaluations of interventions that have failed to address HR & GE issues in their design can consider if the interventions should have paid closer attention to these areas and how this could have been done.

- For evaluations with a **medium and low level** of resources, the focus on data interpretation is more likely to be on whether capacity development of rights holders and duty bearers has led to a sustainable increase in capacity or whether there have been changes in attitudes, behaviours, institutions and legal frameworks and whether this is likely to lead to an improvement in the rights situation of women and individuals/groups who are marginalized and/or discriminated against. These evaluations may also look at whether an enabling environment for the improvement of the HR & GE situation has been created with the support of the intervention. Finally, as in the analysis above, understanding the factors facilitating or hindering changes is critical to a more profound analysis.

136 In order to distinguish the “changes that have taken place in the target population over the lifetime of the intervention and impacts that can reasonably be attributed to the effect of the intervention” (UNICEF, p. 58), UNICEF proposes to use a contribution analysis to assess what would have been the condition of the target population if the intervention had not taken place. See Michael Bamberger and Marco Segone, ‘How to Design and Manage Equity Focused Evaluations’, UNICEF, 2011, pp. 58-61, <mymande.org/sites/default/files/EWP5_Equity_focused_evaluations.pdf>.
7.2.1. Validation

223. When evaluators have gathered their information and prepared tentative findings, it is good practice to validate these findings through workshops with different groups, to increase their accuracy and reliability and to enhance the sense of ownership of the data and process with all stakeholders. The design may include reporting back key findings to separated or mixed (homogeneous or heterogeneous) groups of stakeholders, to programme implementers, and to external experts. The information can be presented for validation, for deepening the analysis, and for eliciting potential conclusions and recommendations.

224. The selection of participants should refer back to the stakeholder analysis, including special attention to women and individuals/groups who are marginalized and/or discriminated against, who can normally be left out of discussions due to multiple kinds of constraints. To adequately respond to HR & GE, the workshop needs to follow the lines already adopted in the evaluation process: being as inclusive as possible, creating an adequate and safe space for reflection, and generating active, free and meaningful participation.

225. At this point, stakeholders will have a chance to understand how the information they have provided has been used, which is in line with the principles of accountability and transparency. Moreover, it is a chance for stakeholders to correct inaccuracies, to ask questions and clarify points of view. For the evaluators, it is an opportunity to explain how they have dealt with conflicting perspectives encountered during the process, and how they have made sure to integrate the different sides of the story.

226. Conducting the final workshop is an important element of validation of the evaluation results. It adds credibility to the process and enhances the likelihood that stakeholders will use the evaluation results later on. The conclusions of the workshop will be an asset to support the evaluators during the report-writing stage. However, it is important to highlight that this process does not entail looking for agreement and possibly compromise. The evaluated group should have the right to respond formally to the evaluation recommendations (through management response) but they should not interfere with the drafting of the recommendations.

7.3. Evaluation report

227. The UNEG Standards for Evaluation in the UN System include overall HR & GE guidance on the drafting of the evaluation report:

The evaluation report should indicate the extent to which gender issues and relevant human rights considerations were incorporated.

---

137 This workshop is usually named ‘debriefing workshop’ or ‘validation workshop’.
138 <www.uneval.org/document/detail/22>
228. The definition of this standard provides details as to what should be included in the evaluation report. The document should specify:

- How gender issues were implemented as a cross-cutting theme in programming, and if the subject being evaluated gave sufficient attention to promote gender equality and gender sensitivity;
- Whether the subject being evaluated paid attention to effects on women and individuals/groups who are marginalized and/or discriminated against;
- Whether the subject being evaluated was informed by human rights treaties and instruments;
- To what extent the subject being evaluated identified the relevant human rights claims and obligations;
- How gaps were identified in the capacity of rights holders to claim their rights, and of duty bearers to fulfil their obligations, including an analysis of gender and women and individuals/groups who are marginalized and/or discriminated against, and how the design and implementation of the subject being evaluated addressed these gaps;
- How the subject being evaluated monitored and viewed results within this rights framework.

229. All of these elements have been discussed throughout this Guidance, and the issues listed in the UNEG Standards are a useful reminder of the key HR & GE areas that need to be covered. The extent to which they are elaborated on in the report will depend on the attention they have received during the evaluation process and in the intervention evaluated. Where there is a low level of resources invested in analysing the promotion of HR & GE, the evaluation report should clearly indicate the rationale for this choice. For example, a real-time evaluation of an emergency situation which lasts only two weeks and with limited access to the affected population may only be in a position to highlight issues related to protection and gender equality, but not undertake a full HR & GE analysis.

230. According to UNEG’s guidance, a specific section on HR & GE should be included at the end of the report. However, an alternative for HR & GE responsive evaluations would be to highlight the implications for HR & GE under each section of the evaluation report, as described in Table 16.

---

139 It should be noted that UNEG is currently in the process of reviewing the Norms and Standards, and a preliminary analysis has already identified human rights and gender equality as one of the focus areas for the review.
231. A good evaluation report needs to make sure that the information provided by participants during the evaluation process, including the final workshop, is duly captured with balanced perspectives and fair representation of different points of view. Findings and recommendations need to be formulated in detail, identifying to whom the recommendations are addressed and proposing concrete action points. For evaluations of interventions where the main focus is on promoting HR & GE, most recommendations will focus on human rights and gender equality. For evaluations of other interventions, it is important that evaluators integrate HR & GE throughout the evaluation process, including in the formulation of recommendations. The recommendations should clearly specify which evaluation stakeholder they are addressed to. This will facilitate follow-up to recommendations through a management response. The evaluation report is the most important resource for the evaluator to reassert the importance of adequately addressing HR & GE.

<table>
<thead>
<tr>
<th>Table 16. Content and Standards for Evaluation Report</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UNEU standards for report content</strong></td>
</tr>
<tr>
<td><strong>Object of the evaluation</strong>: a description of the intervention being evaluated, including:</td>
</tr>
<tr>
<td>• its logic model and results chain;</td>
</tr>
<tr>
<td>• its scale and complexity (number of components, geographic context, total resources);</td>
</tr>
<tr>
<td>• stakeholders involved;</td>
</tr>
<tr>
<td>• implementation status.</td>
</tr>
<tr>
<td><strong>Evaluation purpose, objective(s) and scope</strong>: a general and clear description of the evaluation, including:</td>
</tr>
<tr>
<td>• purpose of the evaluation (rationale behind the need for the evaluation, evaluation users, type of information needed and how it will be used);</td>
</tr>
<tr>
<td>• objectives and scope (evaluation questions, coverage, justification for what was not covered);</td>
</tr>
<tr>
<td>• evaluation criteria;</td>
</tr>
<tr>
<td>• gender and human rights in the evaluation scope.</td>
</tr>
<tr>
<td>UNEG standards for report content</td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
</tbody>
</table>
| **Evaluation methodology:** description of the methodology applied to the evaluation that clearly explains how the evaluation was specifically designed to address the evaluation criteria, yield answers to the evaluation questions and achieve evaluation purposes, including:  
  - data collection methods and analysis, the rationale for selecting them, and their limitations;  
  - data sources (rationale for their selection, limitations, how the mix of data sources was used to obtain a diversity of perspectives, ensure data accuracy and overcome data limits);  
  - sampling frame (area and population represented, rationale for selection, mechanics of selection, limitations of the sample);  
  - stakeholder’s consultation process;  
  - methods employed to answer evaluation questions and to address gender and human rights;  
  - measures taken to ensure data quality, reliability and validity of data collection tools (e.g. interview protocols, observation tools, etc.) | The report should describe:  
  - data collection methods designed to address HR & GE issues;  
  - diversity of perspectives in data sources and processes to guarantee protection of subjects and respect for confidentiality;  
  - how the sampling frame addressed the diversity of stakeholders in the intervention, particularly women and individuals/groups who are marginalized and/or discriminated against;  
  - participatory tools for consultation with stakeholder groups, and the level of inclusion of women and individuals/groups who are marginalized and/or discriminated against in the consultation process;  
  - evaluation questions related to HR & GE;  
  - validation processes responsive to HR & GE. |
| **Findings:** description of evaluation findings according to the evaluation criteria and questions, including:  
  - systematic and appropriate analysis and interpretation of the data;  
  - specific findings addressing each criterion and question posed by the evaluation;  
  - evidence of findings;  
  - gaps and limitations in the data and/or unanticipated findings;  
  - reasons for accomplishments and failures, including constraints to the success of the intervention. | The report should describe:  
  - analysis and interpretation of data on HR & GE;  
  - specific findings on HR & GE-related criteria and questions;  
  - evidence of findings related to HR & GE;  
  - gaps and limitations to addressing HR & GE;  
  - unanticipated effects of the intervention on HR & GE issues;  
  - factors facilitating or hindering success in the area of HR & GE. |
| **Conclusions:** judgements, insights and lessons related to the intervention, including:  
  - identification and/or solutions of important problems or issues pertinent to the prospective decisions and actions of evaluation users;  
  - strengths and weaknesses of the intervention, based on the evidence presented and taking due account of the views of a diverse cross-section of stakeholders. | The report should describe:  
  - insights and lessons regarding HR & GE in the intervention;  
  - identification and/or solutions of HR & GE problems or issues in the intervention;  
  - strengths and weaknesses of the intervention regarding HR & GE;  
  - evidence that conclusions have taken into consideration the perspectives of the intervention’s diversity of stakeholder groups. |
Table 16. Content and Standards for Evaluation Report  (continued)

<table>
<thead>
<tr>
<th>UNEG standards for report content</th>
<th>Implications for HR &amp; GE responsive evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recommendations</strong>: recommendations on the intervention, supported by evidence and conclusions, developed with the involvement of relevant stakeholders, including:</td>
<td>The report should describe:</td>
</tr>
<tr>
<td>• process followed in developing the recommendations including consultation with stakeholders;</td>
<td>• how the process for developing recommendations has involved the intervention’s diversity of stakeholder groups;</td>
</tr>
<tr>
<td>• relevant recommendations to the intervention;</td>
<td>• specific recommendations addressing HR &amp; GE issues;</td>
</tr>
<tr>
<td>• target group for each recommendation;</td>
<td>• target group for HR &amp; GE-related recommendations;</td>
</tr>
<tr>
<td>• actionable recommendations that reflect an understanding of the commissioning organization</td>
<td>• how recommendations on HR &amp; GE reflect understanding of the context, organizations and stakeholders involved in the intervention;</td>
</tr>
<tr>
<td>and potential constraints to follow-up;</td>
<td>• priorities for action to improve the HR &amp; GE dimensions of the intervention or future initiatives in the area.</td>
</tr>
<tr>
<td>• priorities for action.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>All provisions are applicable to HR &amp; GE responsive evaluation reports.</td>
</tr>
</tbody>
</table>

**Gender and human rights**: extent to which the design and implementation of the intervention, the assessment of results and the evaluation process incorporate a gender equality perspective and human rights-based approach, including:

• using gender sensitive and human rights-based language throughout the report, including data disaggregated by sex, age, disability, etc.;
• how the evaluation approach and data collection and analysis methods are gender equality and human rights responsive and appropriate for analysing the gender equality and human rights issues identified in the scope;
• judgement whether the design of the intervention was based on a sound gender analysis and human rights analysis and implementation for results was monitored through gender and human rights frameworks, as well as the actual results on gender equality and human rights;
• findings, conclusions, recommendations and lessons that provide adequate information on gender equality and human rights dimensions.