The Walloon Institute for Evaluation, Forecasting, and Statistics.

Wallonie

Iweps

WORKING PAPER

N° 38 | Integrating cross-cutting dimensions into evaluations: from gender-sensitive evaluation to sustainability-sensitive evaluation.

SUMMARY:

The persistent gender inequalities in all areas of our societies represent injustices that public policies must address. More broadly, in the face of climate change, overexploitation of natural resources, and inequalities between countries and individuals, the analysis of public policies and their evaluation must integrate all the social, economic, and environmental effects they generate.

The purpose of this Working Paper is to highlight frameworks, approaches, tools, and methods that allow for the integration of the cross-cutting dimensions of gender and sustainable development in the evaluation of public policies. It aims to contribute to the adaptation of the discipline and the evaluation community to the major transformations that our society and policies are undergoing. Gender analysis matrices and approaches like ISE4GEMs (Inclusive Systemic Evaluation for Gender Equality, Environments,

and Marginalized Voices present themselves as concrete toolkits that go beyond the rather conceptual level of gender mainstreaming or sustainable development. They are a means to facilitate the transformation of evaluators' practices towards greater consideration of the complexity of public policies. Nevertheless, to promote an effective and lasting transformation of our society towards greater equality, inclusion, and sustainability, the involvement and commitment of all actors are paramount. Through knowledge sharing, critical reflection, and dialogue, which are at the heart of the ISE4GEMs approach, evaluation can initiate or lubricate the gears of a process of mutual learning. This process must be long-term and extend to the entire society to give the attention they deserve to gender issues, social inclusion, and the environment, alongside purely economic issues.

Muriel FONDER (IWEPS)
Sîle O'DORCHAI (IWEPS)

COLOPHON

This work does not reflect the position of IWEPS and only represents the views of its authors.

Table of contents

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

3

Acknowledgements

4

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

1. Introduction

Evaluation, which is by definition normative, consists of passing judgement on the value of a public intervention on the basis of specified standards and criteria. In general, the value of a public intervention is essentially determined by the standard of achievement of the objectives assigned to it.

This normative aspect of evaluation is all the more important when we are interested in the gendered effects of public policies. In our societies, gender inequalities (see Box 1) persist in many areas (employment, wages, poverty, division of time, participation in decision-making, etc.), despite the existence of a whole arsenal of legislation designed to combat them. Faced with such injustices, beyond the integration of a judgement criterion relating to inequalities, the choice of the transformative paradigm, among the various paradigms that co-exist in evaluation, seems the most appropriate. It guides reflection on the entire evaluation process, so that it incorporates the desire for change towards a fairer, more democratic and more inclusive society.

Box 1: Sex, gender and intersectionality

Sex and gender are two concepts that are still often confused. Sex is a biological belonging: every person is assigned a sex at birth on the basis of visible organs. On the other hand, what gives sex a reality are gender norms (cf. Butler, 1990). «*Sex is 'what we have', gender is 'what we are told to do with it' but also 'what we do with it'*» (Amsellem-Mainguy, Gelly and Vuattoux 2017). To give a simple definition, gender is a social construction of relations between the sexes and the roles assigned to men and women in society.

Intersectionality takes us a step further, and refers to the idea that different power relations (based on gender, physical characteristics, ethnic origin, age, religion, etc.) are closely intertwined, and that certain people or groups suffer several forms of domination and discrimination at the same time. This concept was created in 1989 by Kimberley W. Crenshaw, an American Afro-feminist who devoted her academic research to *Black Feminism*. This American feminist movement asserts the particular point of view of African-American women both on feminism in general and on the struggles against segregation between whites and blacks. The meaning of the term has since been broadened, in the 2010s, with the rise of cyber-militancy, and now encompasses all forms of discrimination that may intersect.

In a context of unlimited needs and limited public resources to meet them, it is important to allocate these resources as effectively and efficiently as possible, avoiding any wastage. Hence the importance of properly identifying and analysing the needs to which public intervention claims to want to respond. *Mainstream* approaches that ignore the existing differences between women and men lead to policies that are ill-adapted or calibrated and under-efficient.

The integration of a gender perspective has been an obligation since 2007, when Belgium adopted its law on *gender mainstreaming* (an integrated approach to the gender dimension). The aim of this strategy is to strengthen the equality of women and men in society by integrating the gender dimension into all stages of public policy: development, implementation and evaluation (Fallon 2004). A policy that integrates the gender dimension is one in which the situation of the women and men concerned has been examined comparatively, where any inequalities between the sexes have been identified and which seeks to reduce or eliminate them. Gender *mainstreaming* is therefore a structural approach that covers all phases of the "policy cycle", including the evaluation phase, and is binding on all the players involved. Finally, *gender mainstreaming* is a cross-cutting approach,

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

5

meaning that it applies to all policy areas, "*requiring that the vertical walls between policy sectors be broken down in favour of cross-cutting governance*" (Woodward 2003). In terms of governance, at a time of climate change, overexploitation of natural resources and inequalities between countries and individuals, the analysis and evaluation of public policies must take into account all the social, economic and environmental effects they generate, in other words, opt for evaluations that are sensitive to sustainability (see Box 2).

Box 2: Sustainability and sustainable development

The concept of sustainability is a direct descendant of *sustainable development*. The most widely accepted definition is "*development that meets the needs of the present without compromising the ability of future generations to meet their own needs*" (World Commission on Environment and Development, 1988). The popularity of the term sustainable development is undoubtedly linked to the multiplicity of views surrounding the term, or at least its operationalisation.

In the French language, the term 'durabilité' can lead to misunderstandings, as 'durable' can be understood as being linked to duration and a certain stability. This interpretation is too restrictive for our purposes.

In this chapter, sustainability is defined as the consideration of the social, economic and environmental dimensions (gender, social inclusion, economic development, natural resources, etc.) of an intervention, taking into account the interdependencies between these dimensions and the temporality of the effects of the intervention on them. The question of possible trade-offs between more or less long-term effects and between dimensions is central to sustainability-sensitive evaluation.

How can these cross-cutting dimensions be integrated into the evaluation of public policies? This is the purpose of this *Working Paper*, which is aimed at both sponsors and evaluators to support them in this process. A great deal of work has been done for a long time on the question of gender and its integration into the discipline of evaluation. We will therefore begin by adopting a gender perspective. We will then look at how lenses can be adapted to adopt this multi-dimensional and intersectional perspective on sustainability.

6

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

2. What is gender-sensitive evaluation?

Gender-sensitive evaluation is an approach to evaluation[1] which focuses on the issue of gender. This approach to evaluation is by no means limited to interventions that respond explicitly to a gender-related need or problem, but can be applied to interventions in any field. This integration of gender into evaluation can be more or less pronounced and it is interesting to consider it not in a binary way but as a continuum[2].

Figure 1: Continuum of integration of gender in policy evaluation
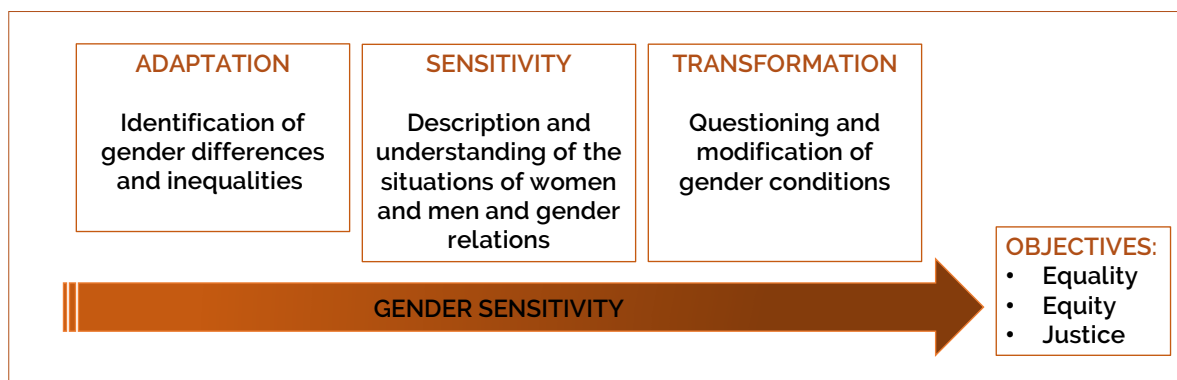
If we exclude gender-blind evaluations, three categories of evaluation can be distinguished on this continuum to clarify the different ways of integrating gender:

- **Adaptation:** The evaluation assesses the intervention by taking into account gender differences and inequalities. This is a relatively cost-effective approach. Identifying gender differences can encourage stakeholders to analyse the sources and consequences of these differences and trigger a process towards greater gender sensitivity. However, this adaptation is often limited to the distinction between women and men in terms of participation in the intervention and results. Both groups are made visible, but without any gender analysis or participation in the evaluation process.

- **Sensitivity:** The evaluation assesses the intervention by analysing the rights, roles, resources, responsibilities and behaviours associated with being female or male, as well as the power relationships between men and women.

- **Transformation:** Evaluation sets out a political and normative agenda for changing gender relations. The values of equality, equity and social justice for the different sexes and, more broadly, for all vulnerable people, are explicit and mobilised within the evaluation. On the one hand, the content of the evaluative research will include an analysis of the beliefs and mechanisms that maintain these unbalanced power relationships and, on the other, the evaluative process will aim to empower vulnerable or marginalised people through their participation it.

---

[1] The term "evaluative approach" refers to the way in which evaluators interact with the parties involved in the evaluation (Ridde and Dagenais, 2009, p. 27). The term "approach" is also used to highlight an angle or objective of the evaluation or its process, such as gender, conflict prediction or the use of evaluation (Befani, 2016, p.14).

[2] Adapted: (1) the IGWG (Interagency Gender Working Group) gender continuum https://www.igwg.org/about-igwg/#continuum (designed to understand gender at programme level, it is here mobilised for the public action instrument of public policy evaluation); (2) the gendered evaluation scale (Bamberger and Mabry, 2020, p. 364).

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

7

In the remainder of this chapter, gender-sensitive evaluation will be understood as any evaluation that goes beyond adaptation and adopts *a minimum* of gender sensitivity. The approach assumes that in order to reduce inequalities between women and men, it is essential to take account of existing power relations. It is based on feminist contributions to the field of public policy evaluation, their focus on the systemic and structural nature of gender inequalities and their demand for greater social justice (Bustelo, 2011; Sielbeck-Bowen et al., 2002). Gender-sensitive evaluation is a key tool for exploring the structural causes of gender inequalities and for determining the differential implications of public policies for women and men (De Waal, 2006).

This gender-sensitive approach can be anchored in different evaluative models and approaches and can accommodate different paradigms. Referring to the "Evaluation Tree"[3] (see Figure 2), a gender-sensitive evaluation in the methods area could be based on a theory-based design that highlights the values and interests of under-represented audiences. A gender-sensitive evaluation in the area of use could adopt a use-focused approach that encourages the involvement of people in the process and thus offers them the opportunity for capacity building. The gender-sensitive appraisal approach, however, has the most affinity with the values branch.

Figure 2: Evaluation tree
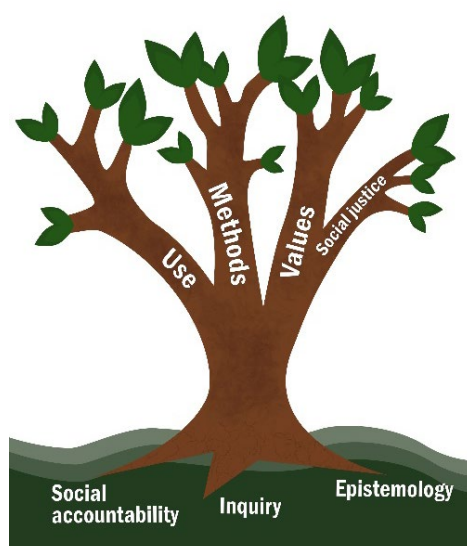
Source: Adapted from Alkin (2012) and Mertens (2018)

Within this field, certain evaluation theorists, such as Ernest R. House (1980, 1990, 2015), Donna M. Mertens (1999, 2010) and Jennifer Greene (1997), focus on the representation of marginalised groups and the issue of social justice. Social justice is the overriding principle that will guide the evaluator's work. This is why Mertens (2018) identifies not three but four branches in evaluative practice: the methods branch, the use branch, the values branch and, finally, the social justice branch. These branches come under the (post-)positivist, pragmatic, constructivist and transformative paradigms respectively. The most gender-sensitive evaluations come under this latter paradigm, whose main philosophical foundations are set out in Box 3.

---

[3] Alkin (2012) and Christie and Alkin (2013) have used the metaphor of the tree to classify different models and approaches to evaluation. Evaluation has its roots in the need for social responsibility, the use of social science research methods and epistemological discussions. The various models and approaches have developed by focusing either on the use of evaluation (use branch), or on the methods to be used (methods branch), or on the evaluative judgement (judgement branch).

8

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

Box 3: The transformative paradigm and the social justice branch (Mertens, 2010, 2016, 2018)

Each paradigm can be defined on the basis of assumptions relating to four areas: ontology, axiology, epistemology and methodology. The transformative paradigm is based on the following assumptions:

- **The ontological hypothesis about the nature of reality**

  This hypothesis recognizes the multidimensional nature of reality. There are as many different perceptions of reality as there are individuals. The evaluator is responsible for making these different perceptions of reality and the factors that influence them (gender, physical characteristics, ethnic origin, socio-economic status, education, religion, etc.) visible, so that the consequences of accepting one version of reality rather than another can be critically examined. This involves encouraging stakeholders to critically re-examine their own representations and assumptions about public intervention and the groups it aims to reach and collecting data directly from them.

- **The axiological hypothesis on the nature of values and ethics**

  This hypothesis reflects an awareness of the pervasiveness of discrimination and the evaluator's responsibility to understand the critical dimensions of diversity in order to challenge societal processes that perpetuate an oppressive *status quo* and thus contribute to the identification of actions that can bring about greater social justice in society.

- **The epistemological hypothesis on the nature of knowledge and the relationship between the knower and what is known (i.e. between the evaluator and the target groups)**

  Knowledge is neither absolute nor relative; it is created in a context of power and privilege. The evaluator needs to develop collaborative relationships that are culturally sensitive and responsive to the needs of different stakeholders in order to create the conditions for the revelation of knowledge from different perspectives and positions. Effective communication strategies are essential to mitigate tensions that may arise due to power differences and the need to address sensitive issues related to discrimination and oppression.

- **The methodological hypothesis on the nature of the systematic survey**

  No particular methodological choice is dictated. The idea is rather to favour the method that will best facilitate (1) the use of the process and results to improve social justice; (2) the identification of the systemic forces that support the *status quo* and those that will enable change to occur; and (3) the recognition of the need for a critical and reflexive relationship between the evaluator and the stakeholders.

However, the term 'social justice' is too restrictive to cover all the difficult issues facing society. Contributing to a transformation of society requires environmental and economic justice to be taken into account (Mertens 2016, p. 157-158).

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

9

The main obstacles to more systematic and widespread gender mainstreaming are undoubtedly:

- Lack of political will: political decisions about what to analyse and evaluate tend to marginalise this topic by limiting gender mainstreaming to evaluations of interventions that specifically aim at gender transformations;

- Limited institutional capacity: the lack of a clear division of responsibilities and insufficient human and financial resources dedicated to gender-sensitive evaluations.

There are, however, two factors which point to a more promising future, and these are discussed in the next two sections. The first concerns the knowledge accumulated and the tools developed, which are now available to the evaluation community to support gender-sensitive evaluations. The second relates to the broad political support given to the sustainable development goals, which include a specific goal relating to gender "Achieve gender equality and empower all women and girls", while at the same time taking a cross-cutting approach to the issue in combination with other goals. The guidelines issued for assessing these sustainable development goals thus reflect a gender perspective.

10

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

3. How to carry out a "gender-sensitive" evaluation/render an evaluation "gender-sensitive"?

Integrating gender into evaluation means adopting a specific approach to the entire evaluation process. In practical terms, this means:

- knowledge of the main dimensions of gender inequality: women's and men's participation in the private and public spheres, women's control over their bodies, practical and strategic gender needs[4], differences in women's and men's time use and unequal access to and control over resources and services;

- be aware of the strong link between gender inequalities and other forms of inequalities in an intersectional perspective (oppressions tend to combine, there is often a confluence of reasons involved, with physical characteristics, ethnic origin, disability, etc. reinforcing gender inequalities);

- revise the entire evaluation process, from preparation to use, to make the evaluation as gender-sensitive as possible.

The figure below, inspired by Espinosa (2013), sets out the essential anchor points for a gender perspective in the evaluation process. The remainder of this section develops these fundamentals in turn for a gender-sensitive evaluation.

Figure 3: Making the entire evaluation process "gender-sensitive"

---
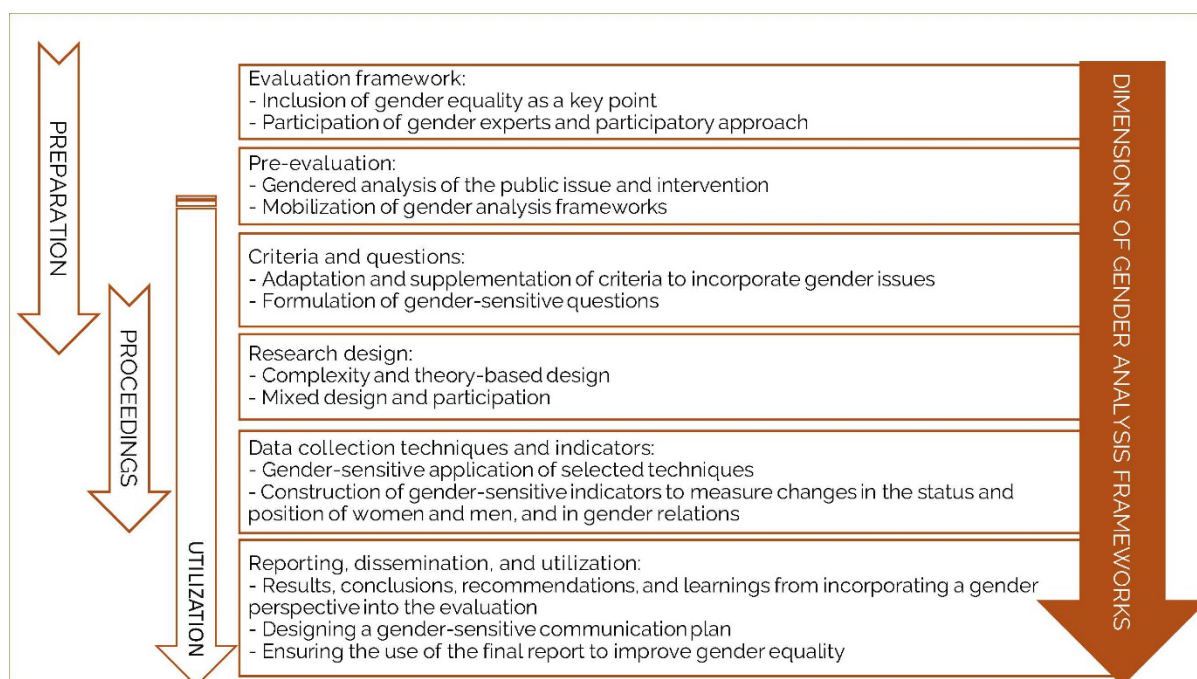
[4] Practical gender needs refer to needs that arise as a result of women's role in society but do not question this role, such as inadequacies in living conditions, health care, the situation on the labour market, etc.
Strategic gender needs refer to the needs that women identify to overcome their subordinate position in society; they are related to the sexual division of labour, power and control, and may include issues such as legal rights, gender-based violence, equal pay, etc. (cf. ILO, 2020, Guidance note 3.1 : Intégrer l'égalité entre les femmes et les hommes dans le suivi et l'évaluation, v.3, https://www.ilo.org/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_746716.pdf).

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

11

3.1. EVALUATION FRAMEWORK

As the sponsor of a gender-sensitive evaluation, it is essential to explicitly mention the "gender-sensitive" nature of the evaluation from the outset. Certain points need to be clarified in terms of what is expected of the evaluation and how these expectations are communicated to potential subcontractors (Espinosa, 2013). The evaluation mandate and terms of reference need to be thought through in terms of a series of different dimensions, all of which are important for guaranteeing the gender sensitivity of the evaluation to be carried out.

3.1.1. The evaluation team

It is important to specify what type of knowledge and experience is expected (general knowledge of gender inequalities, experience in the field of gender analysis, etc.) and whether this is an exclusion criterion or an asset. In some cases, it may be possible to organise a ramp-up period after the contract has been awarded, through participation in awareness-raising and training activities or recourse to specialist external expertise at defined points in the evaluation process. The terms of reference must be precise on these issues.

The gender of the external evaluator or the gender balance of the evaluation team may also be relevant, particularly when the evaluation concerns a sensitive subject (e.g. domestic violence). If the evaluation is to be carried out by a person of a particular gender or by a team with a certain gender balance, the terms of reference must ensure that this expectation is formulated in compliance with employment law.

3.1.2. The thematic content of the evaluation

While evaluations of interventions aimed at achieving gender equality objectives integrate the gender dimension, evaluations of any other intervention do not systematically address the gender issue, even though the intervention could support an oppressive *status quo* or produce gender-differentiated effects. Incorporating gendered evaluation criteria and/or gendered evaluation questions makes it possible to make the request for a gender-sensitive evaluation explicit and concrete.

If a gender analysis of the intervention has not yet been carried out when the mandate and the specifications for the evaluation are drawn up, these documents should explicitly request that the pre-evaluation include this analysis in order to support the identification of the evaluative questions that will form the backbone of the evaluation. Murthy (2016) shows that classic forms of intervention logic, such as the logic model or the theory of change, have been adapted within different bodies to incorporate the gender dimension. Other models have also emerged, such as the matrix of change (cf. *infra*). The intervention logic, whatever form it takes, makes it possible to highlight hypotheses about possible gender effects, or to bring them to the attention of stakeholders for discussion, which feeds into the development of gender-sensitive evaluation questions. Depending on the degree to which gender is integrated (see Figure 1), the questions may focus on the integration of gender into the design of the intervention, whether for the purposes of intervention effectiveness or to meet the specific needs of men and women, or even to question the intervention's contribution to rebalancing gender-based power relations (Chigateri and Saha, 2016).

3.1.3. The mechanism for integrating stakeholders into the evaluation

A major principle of gender-sensitive evaluation is to place women (in general, the vulnerable group in terms of gender-based power relations) and their material realities at the centre of the evaluation. To achieve this, gender-responsive evaluation relies on "**an inclusive, participatory process that respects all stakeholders**" (UN Women, 2015, p. 4). Box 3 details the many benefits of stakeholder

12

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

involvement. However, participation is not a dichotomy, and a number of choices need to be made and should be spelt out in the specifications: the identity of the stakeholders to be included, the stages at which they will participate and their roles/powers[5].

When the objectives of the evaluation were set, stakeholders may already have been involved, in which case it is in the evaluation team's interest to be informed of this so that contacts can be maintained. In all cases, this involvement should be put in place. Once the stakeholders have been identified, it is necessary **to select a limited number of participants for the evaluation** in order to guarantee the effectiveness of the participatory process. The sponsor and/or the evaluation team will have to weigh up the pros and cons between the benefits and risks, such as reduced impartiality, and the additional resources required in terms of cost and time. The involvement of this group can occur throughout the evaluation process, from the identification of what is to be evaluated to the collection, interpretation and use of the data and results. The evaluation team must ensure that each member has a clear understanding of how they will contribute to the evaluation process. It is useful to draw up a reference document setting out the **roles and responsibilities** of each member of the group.

Box 4: The benefits of stakeholder participation in a gender-sensitive evaluation

• **Contributions:** The different perspectives will enrich the design and approach of the evaluation and will in particular help to: identify gender equality and human rights issues and implications that may not have been clearly identified during the programme design and implementation phase; collectively identify and find solutions to the constraints and challenges of the evaluation; facilitate and guide the conduct of the evaluation; and provide realistic and practical contributions to the review and analysis of the evaluation findings, conclusions and recommendations.

• **Use and purpose:** It is the extent to which an evaluation is used and useful that ultimately determines its success. Participatory approaches to evaluation have demonstrated the growing credibility of evaluation results and consequently their increasing use. Involving stakeholders from the outset of the process, in particular to identify the reasons for the evaluation, helps to manage expectations about what the evaluation will be able to deliver. Stakeholder participation is fundamental to the ownership and promotion of the use of evaluation.

• **Capacity building:** Being involved in the evaluation process is in itself a learning experience and can serve to strengthen the capacity of stakeholders through increased exposure to gender equality and human rights issues and gender-sensitive evaluation approaches. This can help to establish clear links between programmes and broader social change objectives; encourage good practice in monitoring and measuring progress on gender equality and human rights; and can be a step in pushing duty bearers (i.e. civil servants or government authorities) to meet commitments on gender equality and human rights.

• **Accountability:** Bringing duty bearers and rights holders together generates a framework of mutual accountability, transparency and application of the fundamental principles that underpin gender equality and human rights.

• **Empowerment:** Involving stakeholders and individual beneficiaries in all stages of the evaluation process and providing them with a space to determine how a genuine process of reflection and evaluation should be undertaken can give them the means to take ownership of the interventions.

Source : ONU Femmes, 2015, pp. 43-44

---

[5] For more information, see the work of Monnier (1992) and Baron and Monnier (2003) on pluralistic and participatory approaches to public policy evaluation.

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

13

It goes without saying that the budgetary and time resources must be in line with the expectations expressed in the terms of reference.

3.2. PRE-EVALUATION

This stage of analysing the intervention is essential for any evaluation. It enables us to understand the context in which a public problem was brought to the fore, the diagnosis that was made, the process by which the various players developed the intervention, and the causal model by which the intervention is supposed to achieve the objectives set. In the case of a gender-sensitive evaluation, this stage will require more or less resources depending on whether or not the intervention has been designed with gender equality in mind. If the intervention is gender-sensitive, an inventory from a gender perspective will have been carried out and certain mechanisms promoting gender equality or the transformation of power relations will have been identified. If the intervention was designed without explicit consideration of the gender issue, then the evaluation exercise must serve to reveal the positive or negative effects of gender relations, whether at the level of the initial situation or at that of the effects of the intervention itself.

Reconstructing the intervention logic is an opportunity for gender-sensitive evaluation, but it is not a panacea (Hivos, 2014), because the intervention logic could also be blind to the gender dimension, particularly for those interventions whose primary objective is not gender equality. In addition to the programming documents for the intervention, the evaluation team will need to draw on theoretical resources[6] for gender analysis, which will be invaluable tools for updating the assumptions in terms of gender on which the causal model is based. For example, the hypothesis of a causal link between the introduction of evening classes in French as a foreign language and the participation of the target populations may raise new questions if a gender perspective is adopted: are learning methods in the form of face-to-face classes the most appropriate for a female or male population? Are the periods of availability the same for men and women? Shouldn't status on the labour market be taken into account when defining learning methods? etc.

A gender analysis framework provides a structure for organising information on gender roles and relations. There are several[7]: the Harvard analytical framework (or gender-role framework), people-centred planning, the Moser framework (or triple-role framework or Moser-Levy framework), Rani Parker's gender analysis matrix (1993), the change matrix (developed by Rao and Kelleher in 2005, then adapted by Batliwala in 2008), the capacity and vulnerability analysis framework, the women's empowerment framework or Longwe method and, finally, the social relations analysis method. These frameworks are practical instruments, initially designed to help their users integrate a gender analysis into their research and/or policy design, but which are also an important support at the policy evaluation level for reconstructing the intervention logic.

The different gender analysis frameworks share certain common features. For example, they all recognise and emphasise the importance of considering not only paid work but also unpaid work (household and care work). They differ, however, in their scope and above all in their understanding of the nature of power and inequality, and therefore in their assumptions about what needs to be analysed and addressed. Some frameworks pursue efficiency objectives, while others pursue *empowerment* or autonomy for women. To quote Kabeer (1994: 303), "*there is an intimate link*

---

[6] As Albarello, Aubin and Van Haeperen (2016) point out, during the pre-evaluation phase, recourse to theoretical resources, in other words, support from currents of thought that have been built up progressively and validated in different scientific disciplines, is necessary to give the evaluation a real foundation. Theory acts as a "third-party distancer", helping to clarify the evaluation questions, which in turn form the backbone of the evaluation process. Reference to theoretical resources thus helps to strengthen and legitimise the subsequent results of the evaluation.

[7] For a detailed description of these different frameworks and keys to choosing the most appropriate one according to the context and the hypotheses analysed, see March et al. (1999), Warren (2007) and Ligero et al. (2014).

14

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

between ways of thinking and ways of acting". Each framework is based on a limited number of factors considered to be the most important in the extremely large set of factors influencing each situation, and can therefore only produce a crude model of the much more complex reality.

The choice of the appropriate framework will depend on the task ahead, the context and resources available, the values and assumptions of the policy designers and evaluators, etc. In practice, frameworks can also be combined so that they can be adapted as closely as possible to the evaluation that requires them.

In the rest of this section, we have chosen to focus on two of these frameworks, the combination of which, in our view, provides a solid basis for identifying the different impacts of interventions on women and men and thus reconstructing the intervention logic of a public policy in a gender-sensitive way. These are the gender analysis matrix (see Figure 4) and the change matrix (see Figure 5).

In the first matrix, the gender analysis matrix, the impact of an intervention is studied in four areas: work, time, resources (taking into account access and control) and socio-cultural factors. In the area of work, the aim is to identify changes in tasks, the level of skills required and productive capacity. In the area of time, changes in the duration of different tasks are studied. The category of resources refers to changes in access to resources as a result of the intervention, as well as to changes in the extent of control over resources in each group analysed. Finally, socio-cultural factors refer to changes in gender roles or status as a result of the intervention, or changes in the wider social life of the target groups.

Figure 4: Gender analysis matrix

	Work	Time	Resources	Culture
Women				
Men				
Household				
Community				

Source: Parker (1993)

The gender analysis matrix can be used during the pre-evaluation phase to support the development of the intervention logic, whatever the future research design of the evaluation. It is often used beyond the pre-evaluation stage as a support for a theory-based participatory evaluation, whether to assess the potential impact, or subsequently the actual impact, of the intervention on gender relations (cf. *infra*).

It is based on a technique for identifying and analysing differences between the sexes, which is part of a participatory approach. The matrix is completed (and regularly updated to progressively complete the expected and unexpected effects) by a group within the community targeted by the intervention, which preferably includes as many women as men. This framework adapts fairly easily to the lack of quantitative data on gender roles.

As a tool for transformative evaluation (see Figure 1), its use is intended to initiate a process of analysis and evaluation by the very members of the community (targeted by the intervention to be evaluated), who are encouraged to identify and constructively question their representations of gender roles. The technical expertise of outsiders is not seen as essential, but at most as facilitating the gender

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

15

analysis. The transformation envisaged must emanate from and be driven by the stakeholders most directly concerned.

The second matrix, that of change[8], is cited in academic work and used by a wide range of organisations of all sizes, as well as by women, local groups and communities around the world. It was developed by *Gender at Work*, an international feminist knowledge network working to end discrimination against women and build cultures of inclusion. The top two quadrants refer to the individual. On the right are changes in observable individual conditions, such as increased resources, access to health and education, etc. On the left are individual awareness and capacities - knowledge, skills, political awareness and commitment to moving towards equality. The lower part refers to the level of society, of the system. The right-hand quadrant refers to the formal rules set out in constitutions, laws and policies. The left-hand quadrant covers all the informal discriminatory norms and structures that maintain inequality.

Figure 5: Matrix of change
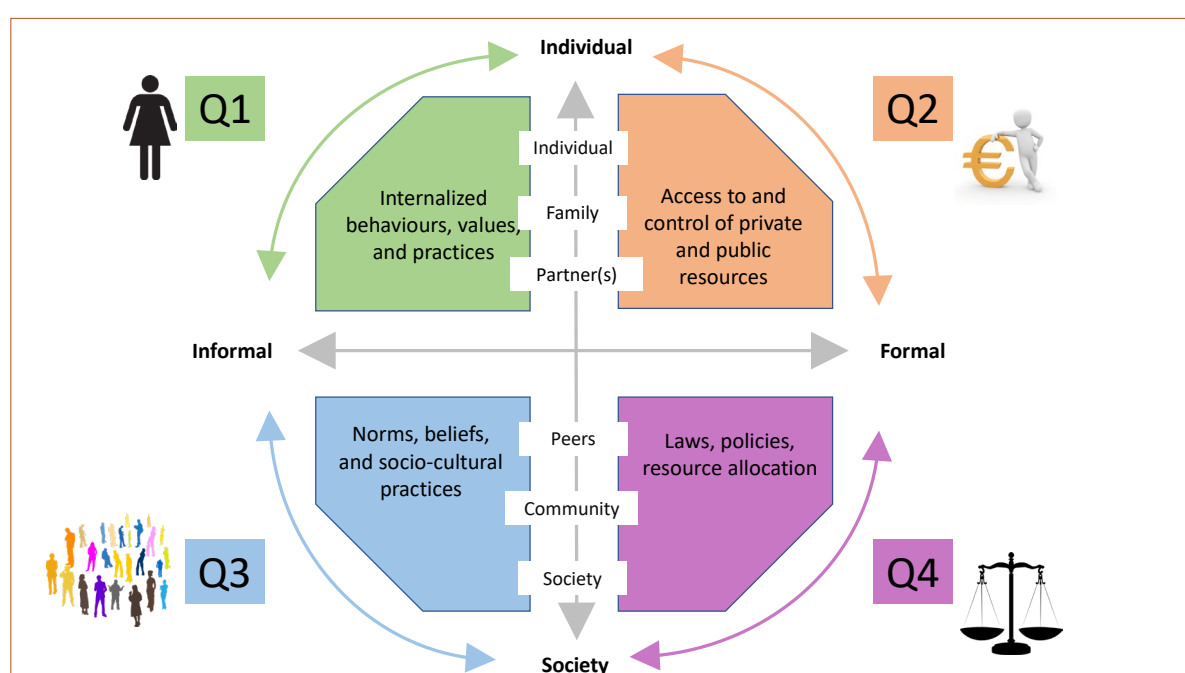
Source: Rao and Kelleher (2005, adapted by Batliwala (2008)

By making visible the different dimensions of gender (in)equality and changes in power relations between the sexes on two axes varying from the individual to society and from the formal to the informal sphere, this matrix adds levels of comparative analysis to the gender analysis matrix, which addresses changes in power relations in four key areas of gender inequality in our societies.

It seems to us that by combining these two analytical frameworks, gender inequality and power relations are analysed in all their facets. Indeed, within each quadrant of the matrix of change, they are studied in the four key areas of work, time, resources and socio-cultural factors.

As practical tools, frameworks such as those detailed above allow certain parts of the intervention logic to be developed, whether in terms of reaching target audiences or various intended or unintended outcomes. As in any evaluation, resources are limited and choices will have to be made as to which aspects of the intervention are most in need of information and/or which are of most

---

[8] Originally designed by Rao and Kelleher (2005).

value to the sponsor and stakeholders. These choices will be reflected in the evaluation criteria and questions.

3.3.   CRITERIA AND QUESTIONS

Gender-sensitive evaluation involves adapting the evaluation criteria to take account of different dimensions of gender, such as the gendered division of roles and work, women's and men's participation in the private and public spheres, control over women's bodies, practical and strategic gender needs, women's and men's use of time, access to and control over resources by both sexes, and so on. As Sierra Leguina (2000: 97[9]) explains: "*Gender-sensitive evaluation requires a reinterpretation of the criteria to enable an analysis that is less action-centred and more focused on people and processes*". More specifically, the criteria must take into account the socio-cultural aspects and the economic and political structures that generate gender inequalities.

There is no consensus in the literature[10] on how to make evaluation criteria gender-sensitive. Two possible approaches are outlined below.

- Espinosa (2013) uses the OECD/DAC criteria to take gender into account. The following version has been adapted to take account of the 2019 revision of the criteria:
    - **Relevance:** this criterion refers to the extent to which the objectives of the intervention are adjusted to take account of the different problems and needs of women and men. This criterion also asks whether the methodology adopted by the intervention helps women to perceive and overcome the constraints and limits imposed on them;
    - **Consistency:** this criterion refers to the compatibility of the intervention with other interventions, and in particular with the gender equality standards and conventions to which the country or institution subscribes;
    - **Effectiveness:** this criterion refers to the extent to which the intervention has achieved its objectives, particularly in terms of the benefits obtained by women and men, or has produced results from the point of view of gender equality, without reference to the costs incurred to obtain them;
    - **Efficiency:** this criterion refers to the extent to which the results in terms of gender equality are achieved at a reasonable cost; it also questions whether the benefits have an equivalent cost for women and men, and whether this cost is distributed equitably;
    - **Impact:** this criterion refers to the contribution of the intervention to a broader gender equality policy, to sectoral equality objectives and to progress towards equality in the long term. The contribution may be direct or indirect and concern expected or unexpected effects;
    - **Viability/Sustainability** [11]**:** this criterion refers to the long-term nature of the added value of the intervention in terms of gender equality. Attention is paid to possible trade-offs between the different dimensions of sustainable development and to the capacity of the systems needed to ensure the continuity of net benefits over time. This is linked to

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

17

the inclusion of strategic gender needs in the intervention and to its appropriation by women and men.

- De Waal (2006) suggests combining the traditional evaluation criteria with the five objectives of *gender mainstreaming*:
    - **Parity,** which refers to the number of women and men participating in or benefiting from a project or action;
    - **Equality (formal)**, which analyses whether women have the same opportunities as men;
    - **Equity (equality of outcome or substantive equality)** which refers to equivalence in the impact on the lives of women and men, recognising their different needs, preferences and interests. This criterion recognises that achieving equality of outcome may require different treatment of women and men;
    - **Empowerment**, i.e. the degree to which women are aware of their subordinate position and the extent to which they are able to counter it;
    - **Transformation**, which refers to the inclusion of gender equality on the political agenda and the incorporation of the gender approach in all policies and programmes, as well as in administrative and financial activities.

The choice of evaluation criteria makes explicit the dimensions of the intervention which will be analysed and whose results will serve as a basis for judging the value of the intervention. Not all evaluations are based on the definition of evaluation criteria, but the evaluation questions themselves contain these criteria.

The evaluation questions, if they are to contribute to a gender-sensitive evaluation, must take account of gender inequalities and start from the assumption that the effects of any public policy are likely to be different for women and men. The evaluation team must reach a consensus among the stakeholders on the questions to be included in the evaluation. These questions can be structured around the chosen evaluation criteria. UN Women (2015) considers that, in general, three to five key questions linked to each of the chosen criteria allow for a more focused evaluation.

The table below lists evaluation questions for each OECD-DAC criterion that adopt a gender perspective.

Table 1: Applying the criteria in a gender-sensitive way

Criteria	Standard questions to ask in order to apply the criterion in a gender-sensitive way
Relevance	Has the intervention been designed to meet the needs and priorities of all genders? If so, how?
	To what extent does the design of the intervention reflect the rights of people of all genders and include feedback/views from a diverse range of local stakeholders, including marginalised groups?
	Does the intervention meet the practical and strategic needs of all genders?
Coherence	To what extent are the design, implementation and outcomes of the intervention consistent with international laws and commitments to gender equality and rights, including the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW), the Beijing Declaration and Platform for Action, the Programme of Action of the International Conference on Population and Development and the 2030 Agenda?

18

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

	To what extent does the intervention support national legislation and initiatives to improve gender equality and human rights? What lessons can be learned?
Effectiveness	Did the intervention achieve its objectives and expected results in such a way as to contribute to gender equality? If so, how?
	Were there different results for different people? If so, how and why? Were different approaches needed to reach people of different genders? Was there sufficient monitoring and analysis of incremental effects? Was the intervention adjusted to address any concerns and maximise effectiveness?
	Was the theory of change and results framework nourished by gender equality, political economy and human rights analysis? If so, to what extent?
	To what extent and why is effectiveness different for people of different genders?
Efficiency	Have the various resources been allocated in such a way as to take gender equality into account? If so, how were they allocated? Was the differential allocation of resources appropriate?
	Do the investment costs per person targeted meet the differentiated needs of people of different genders?
Impact	Were there equal impacts for different genders or were there gender-related differences in participation, experience and impacts? If so, why did these differential impacts occur?
	To what extent do gender-related impacts intersect with other social barriers, including race/ethnicity, disability, age and sexual orientation to contribute to different experiences and outcomes?
	How have gendered norms and barriers within the wider political, economic, religious, legislative and socio-cultural environment impacted on outcomes?
	To what extent have the impacts contributed to equal power relations between people of different genders and to changing social norms and systems?
Viability/ Sustainability	Has the intervention contributed to greater gender equality within wider legal, political, economic and social systems? If so, how and to what extent? Has it led to sustainable changes in social norms that are detrimental to people of all or certain genders? If not, why not?
	Will the achievements in terms of gender equality persist after the end of the intervention? Have the processes helped to sustain these benefits? Have mechanisms been put in place to support the achievement of gender equality in the longer term?

Source: OECD (2021:33)

3.4. RESEARCH DESIGNS

Gender-sensitive evaluation is an exercise in which gender sensitivity must permeate every stage of the evaluation process, including the design of the evaluative research. Rather than a new type of design, it is a question of developing a design that is both technically and contextually appropriate, i.e. that takes account of the evaluative context, the evaluative questions, the potential designs and the nature of the intervention. There is therefore no specific research design for gender-sensitive evaluations, nor is there a standard design that would be suitable for any gender-sensitive evaluation. However, two characteristic elements will generally influence the research design. Firstly, the (usually) complex nature of the intervention and its evaluation, which will lead the evaluation team to

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

19

consider a theory-based evaluation approach. Secondly, the level of gender integration in the evaluation, because the more gender-sensitive the evaluation, the more a mixed methodology with a broad input of qualitative and participatory designs will be required.

3.4.1. *Integrating complexity through theory-based design*

It is customary to distinguish between simple, complicated and complex interventions (Glouberman and Zimmerman, 2002, cited by Rogers, 2008, p. 31). Complexity comes from four essential sources: uncertainty, emergence, feedback[12] and intersectionality (see Box 1). Gender-sensitive evaluations, whether of interventions with a gender objective or of any other interventions that influence gender equality through their implementation or effects, are faced with these complexities. When analysing the intervention, it is essential to understand the different domains and levels of complexity in order to develop a relevant and realistic evaluation approach and research design. Bamberger, Vaessen and Raimondo (2016) take a very broad view of complexity, covering both the complexity of the intervention seen as a system and the complexity of the evaluation. They identify five dimensions of complexity[13], all of which are interconnected. For each of them, we present an example that illustrates the increased difficulty in terms of evaluation:

- The nature of the intervention
  Example: this dimension covers, for example, the objectives of the intervention, which may be multiple and may or may not include a gender equality objective.
- Institutions and stakeholders
  Example: Certain parties involved in the design, implementation or evaluation of the intervention may consider that the intervention has no influence on gender equality. Stakeholders may also not share the same vision of the mechanisms by which the intervention will work. Finally, stakeholders may identify unintended effects.
- Causality and change
  Example: The process of change expected from an intervention in terms of gender equality is often long, changes over time and is subject to feedback effects, which makes it very difficult to assess certain causal issues. Have all the effects been identified?
- Integrating the nature of the system
  Example: An intervention is part of a cultural context with its own norms and beliefs.
- Evaluation
  Example: How can the values of the evaluation team and stakeholders that influence the characteristics of the evaluation be made explicit?

Experimental and quasi-experimental designs, used in isolation, are insufficient to capture the complex links between resources and activities, outputs and outcomes, and context and structures. These designs can only address a limited form of complexity. Theory-based evaluation approaches are therefore often used to assess complex societal change interventions (Fulbright-Anderson, Kubisch and Connell, 1998) where the effects are uncertain and may emerge at different levels. Evaluation should make it possible to check whether the right conditions are in place to achieve the expected impact (Kalpazidou Schmidt & Cacace, 2017) and that the process of producing the expected effect is actually taking place. Theory-based evaluation sees intervention precisely as "a causal chain running from the initial intervention to the final impact, the process of which is not linear, but unstable and beyond the control of the public decision-maker" (Devaux-Spatarakis, 2014, p.54). Theory-based evaluations therefore go beyond purely counterfactual causality in an attempt to

---

[12] Uncertainty relates to the lack of control over elements that affect the intervention and its effects. Emergence concerns what actually happens (whether it is something planned or not) or what could happen. Feedback can take various forms: reinforcement, resistance, repulsion, counterpoint, etc.

[13] For a more detailed presentation, see Bamberger, Vaessen and Raimondo (2016, pp.12-22).

20

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

reconstruct impact mechanisms in detail, involving all stakeholders. Interventions are based on explicit hypotheses formulated by stakeholders that enable us to understand what is happening in the 'black box' and why, what is working or not working, for whom, in what context and, if the expected effects are not being achieved, what needs to be adjusted (Chen, 1990; Pawson and Tilley, 1997; Barnes, Matka and Sullivan, 2003). Theory-based evaluations therefore seek to elucidate how and under what conditions an intervention leads to the expected and unexpected results observed (Döring & Bortz, 2016). They explore not only whether the intervention works, but also how, for whom and in what context (Van Belle, Marchal, Dubourg & Kegels, 2010).

3.4.2. Integrating gender through mixed and participatory designs

The higher the level of gender sensitivity of the evaluation (see Figure 1), the broader the objectives of the evaluation, from observation to explanation, from understanding to transformation of gender relations, including through the evaluation process.

To achieve the objective of understanding complex interventions, quantitative methods based on administrative or survey data are often insufficient. Indeed, mixed methods (combining quantitative and qualitative data and methods) are widely recognised as more appropriate for gender-sensitive assessments "to combine an understanding of the lived experiences of women and men in different types of households, communities and economic activities (qualitative assessment) with an estimation of the significance and representativeness of the observations and the statistical significance of the observed differences (quantitative assessment)" (Bamberger, Segone and Tateossian 2017: 25). Mixed designs therefore rely on the use of a variety of tools, both quantitative (questionnaire surveys, administrative registration data, econometrics, etc.) and qualitative (individual interviews, *focus groups*, direct observation, case studies, etc.).

To achieve the objective of transforming gender relations, in particular through the evaluation process, a mixed design combining quantitative and qualitative methods is still not sufficient in itself. The evaluation process also needs to include a participatory dimension. By involving a wide range of stakeholders (the target audience(s), decision-makers, those responsible for implementation, as well as other players both for and against the intervention), participatory methods make it possible to improve the quality of the evaluation by taking into account a wide range of viewpoints, both in the choice of priority evaluation questions and during the data analysis phase, and then in the drafting of recommendations. Participatory designs support the gathering of complex and rich data and the diversity of the information collected (by supplementing administrative and survey data with more subjective data in the form of opinions, perceptions, representations, etc. of all the parties involved in the process). The participation of women should be envisaged as early as possible in the evaluation process, in order to initiate a process of awareness-raising and learning through the exchange of information and the questioning that takes place during the evaluation. It is this process that is likely to empower the participants and promote change in the more or less long term and through different channels in favour of gender equity.

By way of example, let's take the combination of frameworks proposed above (see Section 3.2), the gender analysis matrix and the matrix of change. They are constructed in several phases. During the pre-assessment phase, an initial fill-in is carried out to gather all the available information on the intervention and its context and, depending on the extent to which the intervention has been implemented, any initial reflections or feedback from the field. Once the matrix has been filled in by the members of the group concerned by the intervention, it can be used to reflect on the desirability

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

21

of the effects identified, their conformity with the objectives of the intervention, the factors that hinder or support the desirable effects, the effects of the intervention on people/groups not directly targeted and/or more vulnerable targeted people, unexpected results, etc. The various stakeholders can draw lessons and conclusions that they will apply from that moment on in their practice. This framework can thus help to anticipate possible resistance and encourage consideration of the support to be offered to people at risk. In the course of an evaluation, these matrices can be reused and updated to improve both understanding of the intervention and the ability of stakeholders to make their voices and arguments heard, develop proposals, and so on. These frameworks are flexible and designed to adapt to changes over time, including unexpected ones. Finally, in the final phase of the evaluation, after filling in the matrix boxes with the (last) changes made to the intervention, the matrices allow the group members to qualify the results/effects identified: a positive sign (+) if the result is consistent with the objectives, a negative sign (-) if the result is contrary to the objectives and a question mark (?) if they don't know. These signs are not intended to be added together to determine the net effect of the intervention. This would oversimplify the picture of a complex reality and distort the mix of positive and negative effects of any intervention.

In addition to the advantages mentioned, the main limitation undoubtedly lies in the fact that, as participatory tools, these matrices certainly allow the women's point of view to be taken into account, but require careful application so as not to create expectations that cannot be met by the evaluation work.

3.5. DATA COLLECTION TECHNIQUES AND INDICATORS

As at the other stages, integrating gender requires certain points of attention in terms of data collection techniques. In particular

- Women and men must be included in the sample or population studied, and potential obstacles to women's participation must be overcome;

- The different time availability of women and men and their different needs and interests must be taken into account;

- An evaluation team with the skills to work with men, women and diverse groups is needed;

- The team must have the skills to strengthen participation and ownership of the evaluation practice and not reinforce traditional leadership and concentration of power.

"Gender-sensitive indicators" or "gender indicators" specifically measure how the intervention promotes gender equality. This type of indicator makes it possible to monitor how the intervention affects the status and position of women and men and relations between the sexes. Gender-sensitive indicators help to explain manifestations of gender inequality that are often invisible in traditional indicators. For example, with regard to employment, gender indicators include all types of work by women, paid and unpaid, which are often not taken into account in traditional indicators. The box below gives another example in the field of poverty.

22

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

Box 5: At-risk-of-poverty indicator

Traditional research on the subject of poverty relies on the very practical instrument of the "household utility function". Poverty measures are based on household resources, which are assumed to be shared equally between the different members of the household, so that everyone has the same poverty status. The standard at-risk-of-poverty rate is calculated as the percentage of people living in households whose equivalised disposable income is less than 60% of the median disposable income in the country. It is therefore estimated on the assumption that all income is pooled and shared between the members of a household, regardless of their own contribution. According to this approach, a person is poor if they belong to a poor household, regardless of their personal income. Basing the measure of a person's at-risk-of-poverty rate on the income of the household to which they belong is problematic, particularly for women. The assumption that resources are pooled and shared within households masks women's precariousness and leads to the implementation of social policies that do not address their particular precariousness and are therefore ineffective.

The availability of sex-disaggregated statistics in all areas, not just those where the gender dimension is obvious, is a prerequisite for the implementation of gender-sensitive evaluations. In the absence of sex-disaggregated statistics to measure the situation of women and men and their roles in social, economic and political life prior to intervention, assessing progress towards gender equality and prioritising actions to address gender inequality issues are particularly difficult. Additional data collection will therefore be required. Because of its importance in advancing the status of women, the development of sex-disaggregated statistics and gender indicators was identified as a priority in the 1995 Beijing Platform for Action.

Furthermore, even if statistics by sex do exist, it is important to understand how they have been constructed. Beyond the example of the at-risk-of-poverty rate (see Box 5), three conceptual and methodological problems can potentially arise:

- Firstly, "sex-neutral" can be synonymous with "*gender-blind*". Where a law makes no distinction between the sexes, it applies equally to women and men. However, since women and men are in different situations, equal application may disadvantage one of the two groups. Making no distinction therefore offers no guarantee that the result will be neutral and unbiased. Neutrality is not synonymous with equality. These biases apply not only to laws but also to the way in which statistical data is collected. For example, not detailing data in the same way for men and women is potentially a source of bias. A typical example is given by the occupational categories which, in the usual classifications, tend to be much more detailed for occupations in which men are over-represented than for those in which women form a majority. The same problem can be observed for fields of study and sectors of activity.

- Secondly, biases may result from the choice of data measurement level. While household data are particularly interesting from a gender perspective, given that inequalities between women and men in various areas (employment, violence, poverty, etc.) are often closely linked to family situation, they become a problem if the data do not allow for disaggregation by gender. Given that gender analyses require sex-disaggregated statistics, household data is a problem. For example, it is not possible to establish, on the basis of the household budget survey, who spends how much, on what, and what the gendered differences in consumption are. Also, in the SILC survey, ownership of an asset, however individual it may be, such as a mobile phone for example, is a question addressed to the household and not to the individual: does the

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

23

household own a mobile phone? This is absurd and problematic. This is why it is important to work with two levels of measurement, supplementing the data collected at household level with individual data.

- Thirdly, household surveys use reference persons or proxies, a practice that is also potentially a source of gender bias. In both cases, information is collected indirectly, with a real risk of distortion:

  - The reference person is someone who 'speaks' for the whole household. The concept of 'reference person' is a more gender-neutral reformulation of the old concept of 'head of household'. Although this is a positive adaptation, it does not resolve all the underlying methodological problems. To choose the reference person within a household, several surveys use criteria that are not gender-neutral, such as age, having a job or highest income, which results in an over-representation of male respondents in the sample used. On the other hand, the birthday method, which involves selecting the last person in a household to have celebrated a birthday, is a gender-neutral method.

  - A proxy is a member of the household who answers in place of the absent respondent. In order to avoid any bias, it is important to limit their use to answering very concrete questions, concerning objective facts, which are in principle very easy to answer. Proxies should not be used to answer subjective questions about satisfaction, quality, perception, opinions and feelings.

3.6. REPORTING, DISSEMINATION AND USE

The final report of a gender-sensitive evaluation will necessarily reflect the specific nature of an evaluation based on the principle of gender equality, in terms of the themes addressed but also the stakeholder integration mechanism. The latter will be explained so as to demonstrate its potential capacity to strengthen the position of certain players and the way in which it avoids the occurrence of additional discrimination and/or the exacerbation of unequal power relations (UN Women, 2015: 91).

The report begins by presenting a gendered analysis of the problematic situation or the needs at the origin of the intervention, on the one hand, and of the intervention evaluated, on the other. It details the way in which the research design incorporated the gender dimension. It explicitly discusses the ability of the intervention to adequately address gender issues, the sensitivity of the intervention to gender issues and draws conclusions in this regard. It concludes with recommendations on how to improve performance in terms of gender equality. It is also appropriate for the report to highlight broader gender equality lessons and good practices that are relevant beyond the immediate scope of the intervention (ILO, 2020).

Once the report has been completed and shared with stakeholders, the dissemination of results and recommendations is a key moment to encourage learning and transformation. Particular attention must be paid to the follow-up of recommendations. Preparing and organising this follow-up triggers the response of the various managers to the evaluation recommendations. The design of a communication plan, which takes into account the specific communication models and spaces for women and men (more or less formal communication, via certain press or social media, via certain venues such as school playgrounds or sports clubs, etc.), is essential to ensure that the report reaches all those involved and that they can all get hold of it and make it their own. Engaging with the media and networking also contribute to the effectiveness of efforts to (re)shape policy (Kelly, 2015, cited in Mertens, 2018:104).

24   IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

This section has provided an overview of the knowledge that has been accumulated in order to incorporate a gender perspective into the evaluation of public policies. However, the practical application of such evaluations is still often limited to evaluations of interventions that explicitly target a gender objective. Designers and implementers are generally very reluctant to accept an evaluation - a judgement on the value - of their intervention that takes into account its contribution to objectives that are not explicitly assigned to it. So how can gender be incorporated as a cross-cutting dimension into evaluations? Hunt & Brouwers (2003) have already mentioned the importance of building a partnership between international institutions, the funders of projects, those involved in evaluation and the various stakeholders concerned with gender equality. In this respect, the European Union has issued a document supporting the development of evaluations that include gender as a cross-cutting dimension, based on the fact that gender equality and the empowerment of women and girls are fundamental human rights and form part of the founding values of the Treaty on the European Union and the Charter of Fundamental Rights (EU, 2018). The United Nations Evaluation Group has adopted a human rights-based approach and a gender mainstreaming strategy as quality standards for evaluations (UNEG, 2016, p.24). Unfortunately, these various guidelines are part of an evaluation context that focuses on external aid to countries in transition or developing countries, and are rarely used to support gender-sensitive evaluations of internal interventions. The broad political support for the sustainable development goals, which include a specific objective relating to gender, appears to be an opportunity to transform the evaluation systems in our countries and make them more sensitive not only to gender but also to environmental issues. In the next section, we will summarise the developments underway to design evaluations that are sensitive to sustainable development. Readers will be able to appreciate the challenges shared by gender awareness and sustainable development awareness, and the similarities in the responses provided.

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

25

4. How has the gender-sensitive evaluation approach spread to the field of sustainable development?

While the concept of gender is the subject of a widely shared definition (see Box 1), the concept of sustainability used in evaluation has yet to be tested. The Bruntland Report set out the contemporary framework for sustainability, defining it as "*development that meets the needs of the present without compromising the ability of future generations to meet their own needs*" (WCED, 1988), and basing it on the three pillars of environment, economy and society. Evaluation practice did not discover the term sustainability following the introduction of the Sustainable Development Goals framework. Among the evaluation criteria of the DAC-OECD, sustainability has been included from the beginning *(DAC-OECD: Development Assistance Committee of the Organisation for Economic Co-operation and Development).* However, its meaning has evolved to adapt to the current context: from a static conception where it was essentially a question of the financial viability of the intervention to a more dynamic conception that encourages analysis of "potential trade-offs between priorities and the resilience of capacities/systems allowing benefits to continue over time" (OECD-DAC, 2019). Although the discipline of evaluation has already touched on these issues of trade-offs between priorities, it has to be said that current evaluations are systematically carried out at the level of a particular theme or sector, reinforcing a siloed approach. Like gender, sustainability must be a cross-cutting dimension, taken into account in all evaluations and evaluation phases.

Of the three pillars, the economic and social pillars, which are related to the human system, are generally well covered by evaluations, even if the gender dimension or attention to the most vulnerable people are less present. With regard to the environmental pillar, which is related to the natural system, apart from interventions where the priority objective is environmental, evaluations struggle to take it into account: results-based management does not encourage the consideration of objectives not explicitly assigned to the intervention, and the evaluation community has little knowledge of natural environment science methods.

The Agenda 2030 and the seventeen Sustainable Development Goals (SDGs) have prompted the community to reflect on its practices and, in particular, on how to integrate several dimensions into evaluation. While gender, the environment and social inclusion are themes studied in the assessments, what is new is that the SDGs recognise the need to integrate socio-ecological considerations into social interventions. Indeed, the term 'environment' is an umbrella term that covers coupled and even intertwined human and natural systems. Changing ecological systems intersect with social interventions aimed at (re)shaping the lifestyles of populations, particularly the most vulnerable. Hence the importance of seeing the environment as an evaluation stakeholder, on an equal footing with the others, in order to decompartmentalise and open up the evaluation to emerging impacts that would otherwise have remained hidden. The aim here is not to reveal, analyse and change the power relationships between women and men, but between human beings and the other forms of life on our planet, nature and the environment, seen as subjects deserving recognition and respect rather than as objects that human beings can manipulate. To quote Pelizzon and Gagliano (2015): "... *the environment as a marginalized stakeholder, lacking voice and diminished by its non-human status*".

What are the bases for developing a sustainability-sensitive evaluation? On the one hand, pioneering work exists and is growing in popularity and, on the other, the development of international

26

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

frameworks for achieving objectives at the global level has prompted reflection on their evaluation. These have led to a new approach, Inclusive Systemic Evaluation Fo(u)r Gender, Environment and Marginalized voices (ISE4GEMs).

4.1.   PIONEERING WORK

The essential foundation of a sustainability-sensitive evaluation is the joint analysis of the two systems: human and natural. Pioneering work, notably by E. Ostrom (1990, 2009) on assessing the sustainability of socio-ecological systems and by W. Clark (2007) on knowledge systems for sustainable development, provides theoretical frameworks that support this type of evaluation. Rowe (2012) outlines the key elements of implementation:

- Start from the assumption that you are evaluating an intervention that is part of the human and natural systems: it is only after investigation that this assumption can be lifted;

- Consider the key mechanism of connectivity between the two systems: the intervention logic must move beyond the temporal and spatial scales of the human system to also take into account those of the natural system;

- Negotiate with the stakeholders to ensure that the evaluation of the intervention is not limited to the priority objectives assigned to it and includes sustainability objectives, even if the intervention was not designed to do so, including in terms of mobilising resources;

- Ensuring that the evaluation team has the knowledge and skills required for both systems: it is important that experts from the natural system are involved from the moment the evaluation questions are formulated and the research design drawn up;

- Define the scope of the evaluation, taking into account the different geographical and temporal scales of the two systems and the different units of analysis: it will be necessary to choose the most relevant scales and units of analysis;

- Identify and include in the process all the stakeholders of the intervention, including representatives/experts of the natural system, and ensure a balanced representation of the various interests;

- Adopt a participatory process to ensure the credibility of the evaluation in the eyes of the main users: make sure that their questions are top of mind, that they understand and accept a language or methods that are less familiar than those of the social sciences, that the communication is adapted to the different audiences, etc.

4.2.   INCLUSIVE SYSTEMIC EVALUATION FO(U)R GENDER EQUALITY, THE ENVIRONMENT AND MARGINALISED VOICES

The work to develop ISE4GEMs (Inclusive Systemic Evaluation fo(u)r Gender equality, Environment and Marginalized voices) was initiated by the independent evaluation service of the UN Women's Group on the basis of the following observation: the establishment of interconnected objectives between environmental, social and economic dimensions, where the degree to which the objectives are achieved depends on the context and is largely beyond the control of the stakeholders, requires the evaluation work to be adapted. The approach developed draws on systems thinking and the consideration of complexity to organise a participatory evaluative approach that highlights the intersectional links shaping the life of human and ecological systems (Stephens, Lewis and Reddy, 2018). Three dimensions are highlighted:

- **Gender** with a view to gender equality;

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

27

- The **Environment**, to analyse the effects of interventions on human-built environments (cities, refugee camps, parks, etc.), ecological systems (forests, marine ecosystems, etc.) and socio-ecological landscapes important for human well-being (mines, farms, oil fields, dams, etc.);
- **Marginalised voices** to address the unequal power of certain groups of people due to their characteristics (age, sexual orientation, ethnic origin, religion, disability, status, etc.), but also of certain non-human voices such as fauna, flora, culture, etc.

The examination of each of these GEMs dimensions is accompanied by an analysis of their interconnections in order to move away from evaluations in silos. However, the ISE4GEMS approach was conceived in the context of development project evaluation and adopts a specific perspective, that of developmental evaluations (Patton, cited in Stephens & al., 2018, p.50). It is therefore not directly adapted to the evaluation of interventions in our countries and/or to the diversity of evaluation practice.

The ISE4GEMs approach is presented in four phases: (1) Preparation and design (design to be understood as the overall framework of the evaluation and not just its research design); (2) Data collection; (3) Data analysis and writing of the evaluation report; (4) Capacity development. This breakdown, which differs from the one we presented for the gender-sensitive approach (see Figure 3), should not hide the similarity of the approaches and elements highlighted (cf. *infra*). In Figure 6, we propose a model approach for a sustainability-sensitive evaluation that incorporates the essential contributions of the ISE4GEMs approach while being applicable to a broad spectrum of evaluations.

Figure 6: Making the entire evaluation process "sensitive to sustainability"
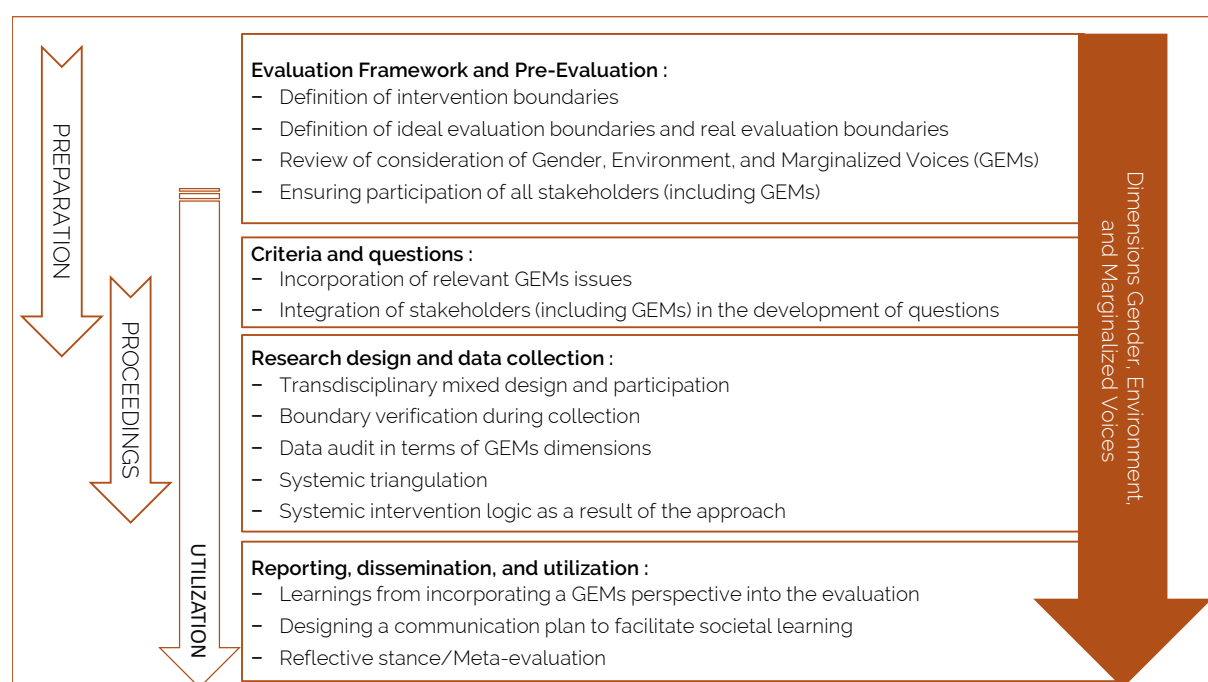

Figure 6 highlights the great similarity between gender-sensitive and sustainability-sensitive evaluations, with each stage of the evaluation process being concerned. On the one hand, taking several dimensions into account requires a holistic vision and an analysis of the complexity of the intervention and its evaluation. Secondly, these approaches are based on the development of a detailed intervention logic as part of an inductive and participatory approach. Finally, particular attention is paid to the use of the whole evaluation exercise, its process and its deliverables, to support change at both the intervention and societal levels. In Figure 6, certain stages have been grouped together, on the one hand, those of establishing the evaluation framework and pre-

28

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

evaluation and, on the other, those of research design and data collection. These groupings are symptomatic of the reluctance to present in a linear fashion a process whose nature is indisputably iterative, *a fortiori* that of participatory evaluations adopting a systemic and complex way of thinking.

In the remainder of this section, we do not present the approach *in extenso* with all its constituent elements (there would be too much repetition compared with the previous sections), but we develop the essential contributions of the ISE4GEMs approach in terms of the preparation, conduct and use of the evaluation.

4.2.1 Preparing and defining the boundaries of the intervention and its evaluation
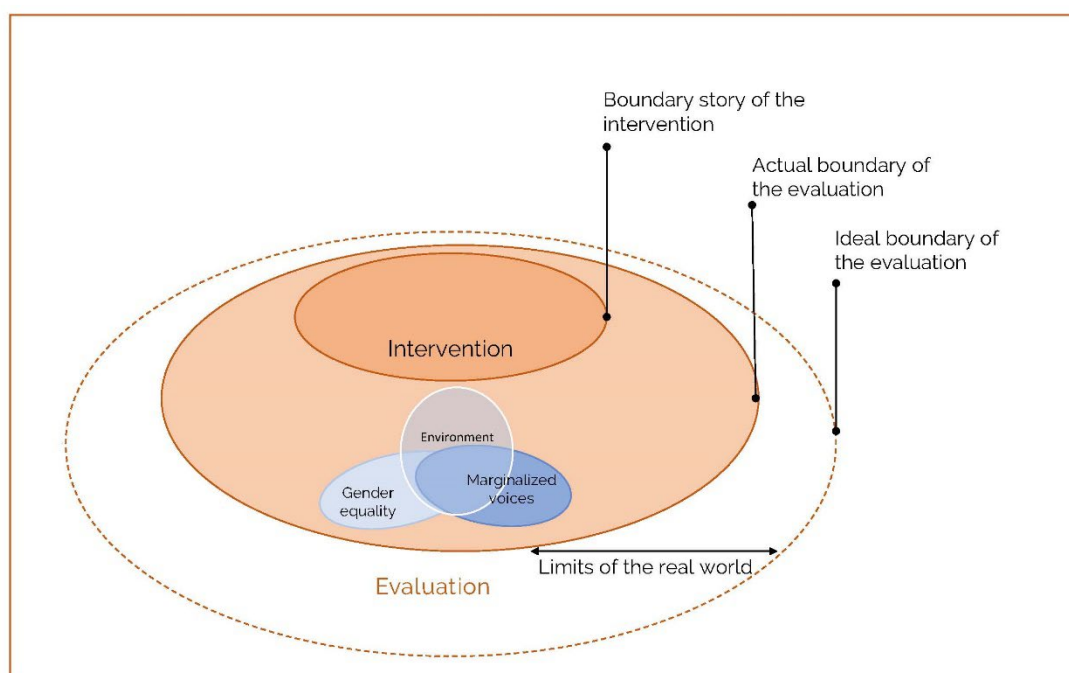
Like gender-sensitive evaluations, the approach is based on a systemic analysis of the intervention, but it also places major emphasis on the analysis of the evaluation system and on the relationships between the two systems. Figure 6 brings together the establishment of the evaluation and pre-evaluation framework, showing the need to connect the two systems. Even if a request for an evaluation is made by a sponsor, the boundaries of the evaluation system will only be effectively defined at the end of a process and these may even be re-examined during the course of the evaluation, to deal with emerging elements of information gathering in the field or new perspectives from stakeholders.

Establishing the boundaries of the intervention requires us to consider the components of the system (people, actions, objectives, etc.), its context (time, geography, culture) and the links with other systems. This static analysis serves as a basis for a more dynamic understanding of the way in which the system was created or has evolved: How was this intervention developed? How was it distinguished from other systems? Who is or isn't part of this system? Who makes the decisions? etc. The form taken by this analysis is a narrative about these boundaries and the changes that may have occurred or been implemented over time. During this first stage, the questioning of the GEMS dimensions will be limited essentially to their inclusion in the development of the intervention, for what reasons, and on their possible connections with the other dimensions. For example, with regard to the environmental dimension, the following questions may guide the analysis:

- Does the intervention identify and address environmental issues? How were they identified? Who identified them? What action has been taken?
- Does the intervention identify and address the relationships between the environment and the other dimensions of gender, the economy and marginalised voices? How were these relationships identified? Who identified them? What action has been taken?

The analysis will initially be based on the strategic or implementation documents underlying the intervention, on an interview with the sponsor and the key players involved in its implementation, on any intervention logic, on monitoring data, etc. This first stage also enables us to identify the various stakeholders, their role(s), their interrelationships and any imbalances in power relations. The involvement of these stakeholders in an initial review of the boundaries of the intervention is an asset (cf. *infra*).

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

29

Figure 7: The boundaries of the different systems

Source: Stephens, Lewis and Reddy, 2018: 61

Figure 7 starts from the intervention boundaries and illustrates the process leading up to the evaluation project.

As mentioned, the GEMs dimensions will not be systematically represented at the level of the intervention; one contribution of the ISE4GEMs approach is to impose their consideration in the evaluation system. The analysis of this system takes place in two stages: the first aims to establish the boundaries of the ideal evaluation and the second, the boundaries of the effective evaluation.

To define the ideal evaluation, the intervention will be examined with regard to the GEMs dimensions. The evaluation team will question any gaps in the story about the boundaries of the intervention and the existence of systems that are nested or interconnected with that of the intervention. It will also, for each of the GEMs dimensions: (1) question the way in which the evaluation will improve accountability and learning with regard to this dimension; (2) identify the elements relating to expected or unexpected results with regard to this dimension; (3) analyse how the evaluation will be used to improve the intervention and promote consideration of this dimension. This questioning helps to identify the dimensions to be included in the evaluation. It is important to note that not all dimensions have the same relevance depending on the intervention, the evaluation period and the context. It is therefore essential to identify for whom one or other of the dimensions is paramount, whether the inclusion of one dimension in the assessment will not hinder the consideration of another due to resource constraints, and the possible implications of these choices in terms of power relations and oppression (Stephens, Lewis and Reddy, 2018, p. 36). The participation of stakeholders representing all the players involved and of experts on the various dimensions should be envisaged from this stage onwards. The feasibility of listening to marginalised human and non-human voices will be analysed and the necessary adaptations studied. Defining the boundaries of the effective evaluation means taking into account the potential vulnerability of certain stakeholders (free access to the evaluation system, risk of harm through participation, etc.), as well as the constraints of the "real world", whether in terms of budget, time, existence or access to data relating to the GEMs dimensions, etc. This dual definition highlights the limits of evaluation work, the distance between

30

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

what would be relevant to evaluate and what is. Recognizing these limits makes it possible to better qualify the results, conclusions and recommendations that are issued (Stephens, Lewis and Reddy, 2018: 58).

4.2.2. Design and data collection process and flexibility

The ISE4GEMs approach promotes the use of mixed and transdisciplinary methods that simultaneously analyse and interpret the interrelationships between the three GEMs dimensions. Structural and inter-connected power relationships and dynamics are examined. A central role is also played by participatory engagement to ensure the plurality of knowledge, viewpoints, perceptions, values and preferences and the ownership of problems and solutions. The production of knowledge is social and shared, and reflective practice is central. The evaluator moves between his roles as expert and facilitator.

The research design and the data collection itself are conceived in such a way as to be able to adapt to the uncertainty of gathering information in the field and to emerging elements during the process: the inclusion of a new stakeholder, the analysis of an unexpected effect observed along the way, insufficient data quality on certain GEMs dimensions, the need to look more closely at the results relating to a sub-group of people, etc. Regular checks are carried out on the boundaries between the ideal evaluation and the actual evaluation. Particular attention is paid to the relationship that has been built up with the stakeholders of the evaluation: does the level of trust allow access to new data or enable certain stakeholders to take a more active role in the evaluation? For example, the ISE4GEMs approach agrees on the value of including in the evaluation team an evaluation officer from within the implementing organisation, or at least someone who is close to the subject being evaluated (Stephens, Lewis and Reddy, 2018: 83), who can contribute his or her knowledge of the context and/or who can acquire evaluation skills during the process that can be used at a later date (cf. *infra*). The evaluation team also reviews its perception of the power relationships between stakeholders and ensures that these are countered. Data audits are also carried out, which may lead to a revision of the design. To allow for these different forms of flexibility, it is essential to provide a buffer in terms of resources, including time.

4.2.3. Process and systemic triangulation

The analysis phase begun during the data collection phase gets a fresh start once all the data has been collected and coded according to the evaluative dimensions and questions and according to the GEMs dimensions. The ISE4GEMs approach recommends identifying strengths and weaknesses in each of the dimensions and then analysing the interconnections between them. For example, a finding of differential access to vocational training for men and women (evaluative dimension of recourse to the intervention and gender dimension) could be connected to an explanation put forward by women relating to the lack of rapid public transport (environmental dimension) which would enable them to avoid being away from home for too long (gender dimension).

To ensure the credibility and ownership of the results of the analysis and the evaluation recommendations, the approach is based on systemic triangulation. This is carried out in three stages: (1) identifying the facts, the evidence about results and changes; (2) allowing the various stakeholders to interpret these results and changes in terms of their values; (3) analysing these results and changes from a systemic perspective, i.e. taking into account their sensitivity to the current definition of the boundaries of the evaluation. Would the results be the same if the temporal boundaries for observing the results were extended, if the geographical boundaries of the intervention were moved to another location, if a change in a system interconnected with that of the intervention took place, etc.? This distancing from the current boundaries of the intervention is crucial in the ISE4GEMs approach, as it

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

31

enables more nuanced and precise conclusions and recommendations to be provided, which take into account the complexity of the change processes as well as a better understanding of the trade-offs to be made.

> Integrating a sustainable development perspective into the evaluation of public policies, like that of gender, means moving away from the comfort of linear 'cause and effect' frameworks and 'traditional' evaluation practices centred on a linear intervention logic to develop results, conclusions and recommendations (Bamberger and Segone, 2011). Theory-based evaluations can meet this challenge. An intervention logic such as the theory of systemic change encourages the mobilisation of a plurality of theories and practices brought together within the same evaluative framework, which is itself evolving (cf. *supra*). Compared with a 'classic' intervention logic (if such a thing exists), the theory of systemic change can disconnect itself from the way in which the intervention under review is supposed to produce change (a vision) and focus on social change, but by analysing it from different perspectives. It can thus serve as a learning tool and be constantly adapted and revised by various sources beyond the evaluative work (Stephens, Lewis and Reddy, 2018: 104).

4.2.4. Use and cultural change

Capacity development is at the heart of the ISE4GEMs approach and is one of its ultimate goals. Although this objective is undoubtedly taking on greater importance in the context of development policies, it retains all its relevance in a developed region such as Wallonia. It is "*empowerment*" in the sense of enablement, "making someone capable of...", rather than emancipation (cf. Bélanger, 2011). These skills consist of making gender inequalities, the vulnerabilities of marginalised groups and environmental issues more visible, forcing them to be taken into account in political decision-making, and enabling the three stakeholders - women, marginalised voices and the environment (as a stakeholder) - to participate in the processes of designing, implementing and evaluating public policies. We can generalise to sustainable development the maxim of gender studies: it is not enough to "*just add women and stir*". Gender and sustainable development necessarily imply cultural sensitivity and consideration of unequal power relations between stakeholders (including the environment).

To this end, the ISE4GEMs approach emphasizes the importance of engaging stakeholders in a dialogue at each phase of the evaluation. The evaluation team must ensure that reflective practice plays a central role in the entire evaluation process, combining the roles of expert and facilitator. The evaluator is thus responsible for making the different perceptions of reality and the factors influencing them visible, in order to critically examine the consequences of adopting one version of reality rather than another. This involves encouraging stakeholders to critically re-examine their own representations and assumptions about the public intervention, the groups it aims to reach and the changes brought about by the intervention. The aim is to collectively understand the interactions, dynamics and models that underpin them. In this way, a process of mutual learning is initiated. Over time, as this approach becomes more widespread, the various stakeholders will acquire better skills to influence the transformation of public policies in a more sustainable direction.

Starting from the empowerment (or enablement) of the stakeholders in the evaluation, it will be a question of extending it to the whole of society in order to develop a genuine general culture of sensitivity to GEMs. In this respect, the emphasis in the ISE4GEMs approach on communication and its many adaptations according to the audiences to be reached is essential if society as a whole is to understand and support change.

32

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

5. Conclusion

The aim of this *Working Paper* is to present frameworks, approaches, tools and methods for integrating the gender and sustainable development dimensions into public policy evaluation. Its aim is to help the discipline and the evaluation community adapt to the major changes taking place in our society and the policies that govern it. Gender and sustainable development are two unavoidable issues facing our changing society. They cut across the board and cannot be confined to certain phases in the life cycle of public policies or to certain policy areas. This cross-cutting approach poses a real challenge in terms of their systematic and widespread integration into evaluation practice.

As far as gender is concerned, it continues to be hampered by a lack of political will. Despite the fact that *gender mainstreaming* is now a legal obligation, political decisions tend to marginalise this subject by limiting the integration of a cross-cutting gender approach to evaluations of interventions that specifically aim to bring about gender transformations. The situation is slightly different for sustainable development: the international framework of the Sustainable Development Goals, to which the Walloon Government subscribes, provides a strong impetus for the practice and widespread integration of sustainable development into evaluation practice.

In addition to the political will, the question of the role of evaluators in this integration of cross-cutting dimensions needs to be raised. Some are reluctant to promote the inclusion of these cross-cutting dimensions when defining the boundaries of an evaluation, unless they are explicitly included in the objectives assigned to the intervention. Furthermore, the evaluation community sometimes still lacks the skills to integrate these dimensions adequately. Solutions are emerging, however, and this *Working Paper* has attempted to share them. Gender analysis frameworks such as the gender and change analysis matrices and the ISE4GEMs approach are concrete, practical toolboxes that go well beyond the fairly conceptual level of *gender mainstreaming* or sustainable development. They are a means of facilitating the transformation of the practice of evaluation officers towards greater consideration of the complexity of public policies.

If we are to promote lasting transformative change in our society towards greater equality, inclusion and sustainability, we need the support and involvement of the whole of society. To promote this effective transformation of reality, all players need to be familiar with factual information, understand the mechanisms underlying interventions, perceive power relationships, etc., all of which they can mobilise in their day-to-day actions. This sharing of knowledge, critical reflection and dialogue is central to the ISE4GEMs approach (as it is to other participatory approaches). In this way, evaluation initiates or oils the wheels of a mutual learning process. This should be a long-term process, extending to the whole of society and developing a genuine general evaluation culture that goes beyond economic issues and is also sensitive to issues of gender, social inclusion and the environment.

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

33

Bibliography

Alkin M.C., (2012) *Evaluation roots : A wider perspective of theorists' views and influences*, Second édition, Sage.

Amsellem-Mainguy, Yaëlle, Maud Gelly et Arthur Vuattoux (2017) « La Santé en action », *Dossier Genre et santé*, n° 441, 2017, pp. 8-45.

Bamberger M., L. Raftree & V. Olazabal (2016) "The role of new information and communication technologies in equity-focused evaluation: Opportunities and challenges", *Evaluation*, 22:2, 228-244.

Bamberger, M. & M. Segone (2011) "How to design and manage equity-focused evaluations", *Evaluation Working Paper*, UNICEF, New York.

Barnes, M., Matka, E. & H. Sullivan (2003) « Evidence, understanding and complexity evaluation in non-linear systems", *Evaluation*, 9:3, 265-284.

Baron Gaëlle et Eric Monnier (2003) « Une approche pluraliste et participative : Coproduire l'évaluation avec la société civile », Informations sociales, n° 110, septembre, 7p.

Bustelo M. (2011) "Last but not least : gender-sensitive evaluation as a forgotten piece of the policymaking process", *mimeo*, 21 p.

Butler J. (1990) *Gender Trouble: Feminism and the Subversion of Identity.* Abingdon: Routledge.

Chen, H. T. (2012) "Theory-driven evaluation: Conceptual framework, application and advancement", in: Strobl, R., Lobermeier O. & W. Heitmeyer (Eds.) *Evaluation von Programmen und Projekten für eine demokratische Kultur*. Fachmedien Wiesbaden : Springer, 17-40.

Clark, W. C. (2007), "Sustainability science: A room of its own", Proceedings of the National Academy of Sciences, 104, 1737–1738.

Cole, G. (1999) "Advancing the development and application of theory-based evaluation in the practice of public health", *American Journal of Evaluation*, 20, 453-470.

Crenshaw, Kimberley W. (1989) "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics", *University of Chicago Legal Forum*, Vol. 1, n°8, 139-167.

De Waal M. (2006) "Evaluating gender mainstreaming in development practices", *Development in Practice*, 16:2, 209-14.

Döring, N. & J. Bortz (2016) *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin : Springer.

Douglas, F. C., Gray, D. A. & E. R. van Teijlingen (2010) "Using a realist approach to evaluate smoking cessation interventions targeting pregnant women and young people", *BMC Health Services Research*, 10:49, 7p.

Espinosa J. (2013) "Moving towards gender-sensitive evaluation? Practices and challenges in international-development evaluation", *Evaluation*, 19:2, 171-182.

EvalPartners (2016), "No one left behind – Evaluating SDGs with an equity-focused and gender-responsive lens", Report of the 2016 New York high-level panel and technical seminar, 63 p.

34

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

European Commission (2018) "Evaluation with gender as a cross-cutting dimension", 6p.

Fallon C. (2014) "Le genre de l'état et ses instruments : L'action publique à l'épreuve du *gender mainstreaming* », *mimeo*, 13p.

Fitz-Gibbon, C. T. & L. L. Morris (1996) "Theory-based evaluation", *Evaluation Practice*, 17:2, 177-184.

Fulbright-Anderson, A., Kubisch, A. & J. Connell (1998) "New approaches to evaluating community initiatives" Washington, DC: Aspen Institute.

Greene, J. (1997). Evaluation as advocacy. Evaluation Practice, 18 (1), 25-35.

House, E. (1980). *Evaluating with validity*. Thousand Oaks, CA : Sage.

House, E. R. (1990). Methodology and justice. In K. A. Sirotnik (Ed.), Evaluation and social justice: Issues in public education. New Directions for Program Evaluation no. 45 (Pp. 23-36). San Francisco : Jossey-Bass

House E. R. (2015) Evaluating : Values, Biases, and Practical Wisdom,

Hunt J. & Brouwers R. (2003), « Review on Gender and Evaluation », Room document n°1, DAC Working Party on Aid Evaluation meeting on 27 – 28 March 2003, 82 p.

Kalpazidou Schmidt, E. & M. Cacace (2017) « Addressing gender inequality in science: The multifaceted challenge of assessing impact", *Research Evaluation*, 26:2, 1-13.

Kalpazidou-Schmidt E. & E. K. Graversen (2020) "Developing a conceptual evaluation framework for gender equality interventions in research and innovation", *Evaluation and Program Planning*, 79, 1-14.

Ligero Lasa J. A., J. Espinosa Fajardo, C. Mormeneo Cortès & M. Bustelo Ruesta (2014) « Making evaluation sensitive to gender and human rights – Different approaches", *research report*, 135p.

Marra, M. (2015) "Cooperating for a more egalitarian society: Complexity theory to evaluate gender equity", *Evaluation*, 21:1, 32-46.

Martínez-Palacios J. & I. Ahedo (2020) « How can critical deliberative theory help to solve the methodological challenges of evaluating from a gender+ perspective?", *Evaluation*, 26:4, 438-455.

Mertens D. M. (1999), "Inclusive Evaluation: Implications of Transformative Theory for Evaluation", American Journal of Evaluation, Vol. 20, No. 1, 1999, pp. 1-14.

Mertens D. M. (2010) « Transformative mixed methods and policy evaluation », *Qualitative Inquiry*, 16:6, 469-474.

Mertens, D. M. (2018) *Mixed methods design in evaluation*. Thousand Oaks (CA) : Sage.

Mertens, D. M. et Wilson, A. T. (2018) *Program Evaluation Theory and Practice : A comprehensive guide*, Second edition, Guilford Press.

Monnier Eric (1992) Évaluations de l'action des pouvoirs publics. Economica, CPE, 258p.

OECD-DAC (2019), *Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use*, DAC Network on Development Evaluation, OECD Publishing, Paris, https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf.

ONU Femmes (2015), « Gestion des évaluations sensibles au genre – Manuel d'évaluation », New-York, 178 p.

Ostrom, E. (1990), « Governing the commons: The evolution of institutions for collective action », Cambridge University Press.

IWEPS Working Paper N°38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

35

Ostrom, E. (2009). A general framework for evaluating sustainability of social– ecological systems. Science, 325, 419–422.

Pawson, R. & N. Tilley (1997) "Realistic evaluation", London : Sage.

Pelizzon, A. & M. Gagliano (2015) "The sentience of plants: Animal rights and the rights of plants intersecting?", *Australian Animal Protection Law Journal*, 11, 5-13.

Reale, E., Nedeva, M., Thomas, D. & E. Primeri (2014) "Evaluation through impact: A different viewpoint", *Fteval Journal for Science and Technology Policy Evaluation*, 39, 36-41.

Rogers P. (2008), "Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions", *Evaluation,* Vol 14(1), 29 – 48.

Rowe A. (2012), "Evaluation of Natural Resource Interventions », American Journal of Evaluation 33(3) 384-394.

Rowe, A. (2019). Sustainability-ready evaluation : A call to action. In G. Julnes (Ed.), Evaluating Sustainability: Evaluative Support for Managing Processes in the Public Interest. New Directions for Evaluation, 162, 29–48.

Sielbeck-Bowen K. A., S. Brisolara, D. Seigart, C. Tischler & E. Whitmore (2002) "Exploring feminist evaluation: The ground from which we rise", *New Directions for Evaluation*, 96 (Special Issue: Feminist Evaluation: Explorations and Experiences), 3-8.

Sierra Leguina, B. (2000) « Criterios para la evaluación con perspectiva de género » [« Criteria for evaluating with a gender perspective »], *Revista Española de Desarrollo y Cooperación*, 6, 95-112.

Stephens A., E. D. Lewis & Sh. Reddy (2018) "Towards an inclusive systemic evaluation for the SDGs: Gender equality, Environments and Marginalized voices (GEMs)", *Evaluation*, 24:2, 220-236.

UE- Ares (2018), « L'évaluation avec le genre comme dimension transversale », Références : Ares(2018)3264752, 20/06/2018.

Van Belle, S. B., Marchal, B., Dubourg, D. & G. Kegels (2010) "How to develop a theory-driven evaluation design? Lessons learned from an adolescent sexual and reproductive health programme in West Africa", *BMC Public Health*, 10:741, 10 p.

Verloo M. & C. Roggeband (1996) "Gender impact assessment: The development of a new instrument in the Netherlands", *Impact Assessment*, 14:1, 3-30.

Warren, H. (2007) "Using Gender-Analysis Frameworks: Theoretical and Practical Reflections, *Gender and Development*, 15:2, 187-198.

Woodward A. (2003) "European gender mainstreaming: Promises and pitfalls of transformative policy", *Review of Policy Research*, 20:1, 65-88.

36

IWEPS Working Paper No.38
Integrating cross-cutting dimensions into assessments:
From gender-sensitive evaluation to sustainability-sensitive evaluation

2024