

MAI 2023

WORKING PAPER

N° 38

Intégrer des dimensions transversales dans les évaluations : de l'évaluation sensible au genre à l'évaluation sensible à la durabilité

RÉSUMÉ

Les inégalités de genre persistantes dans tous les domaines de nos sociétés présentent autant d'injustices que les politiques publiques se doivent de corriger. Plus largement, à l'heure des changements climatiques, de la surexploitation des ressources naturelles et des inégalités entre pays et individus, l'analyse des politiques publiques et leur évaluation doivent intégrer l'ensemble des effets sociaux, économiques, environnementaux qu'elles génèrent.

L'objet de ce *Working Paper* est de mettre en lumière des cadres, approches, outils et méthodes qui permettent d'intégrer les dimensions transversales du genre et du développement durable dans l'évaluation des politiques publiques. Il vise ainsi à contribuer à l'adaptation de la discipline et de la communauté évaluative aux transformations majeures que traversent notre société et nos politiques. Les cadres d'analyse de genre tels que les matrices d'analyse de genre et des changements, ainsi que l'approche ISE4GEMs (*Inclusive Systemic Evaluation for Gender Equality, Environments and Marginalized Voices*

– Évaluation systémique inclusive pour l'égalité de genre, l'environnement et les voix marginalisées), se présentent comme des boîtes à outils concrets dépassant largement le niveau assez conceptuel du *gender mainstreaming* ou encore du développement durable. Ils sont un moyen de faciliter la transformation de la pratique des évaluateurs et évaluatrices vers une plus grande prise en compte de la complexité des politiques publiques. Il n'en reste pas moins que pour favoriser une transformation effective et pérenne de notre société vers davantage d'égalité, d'inclusion et de durabilité, l'adhésion et l'implication de tous les acteurs sont primordiales. Grâce au partage de connaissances, à la réflexion critique et au dialogue qui sont au centre de l'approche ISE4GEMs, l'évaluation peut amorcer ou huiler les rouages d'un processus d'apprentissage mutuel. Ce dernier doit s'inscrire dans la durée et s'étendre à toute la société pour donner l'attention qu'ils méritent aux enjeux du genre, de l'inclusion sociale et de l'environnement, à côté des enjeux purement économiques.

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

COLOPHON

Auteures : **Muriel FONDER** (IWEPS)
Sîle O'DORCHAI (IWEPS)

Édition : **Evelyne ISTACE** (IWEPS)

Ces travaux ne reflètent pas la position de l'IWEPS et n'engagent que leurs auteurs.

Création graphique : **Deligraph**
<http://deligraph.com>

Dépôt légal : D/2023/10158/5

Reproduction autorisée, sauf à des fins commerciales, moyennant mention de la source.

IWEPS

Institut wallon de l'évaluation, de la
prospective et de la statistique

Route de Louvain-La-Neuve, 2
5001 BELGRADE - NAMUR

Tel : 32 (0)81 46 84 11

<http://www.iweps.be>

info@iweps.be

Table des matières

Remerciements.....	4
1. Introduction.....	5
2. Qu'entend-on par évaluation sensible au genre ?	7
3. Comment faire (rendre) une évaluation « sensible au genre » ?	11
3.1. Cadre de l'évaluation.....	12
3.1.1. <i>L'équipe évaluatrice.....</i>	<i>12</i>
3.1.2. <i>Le contenu thématique de l'évaluation</i>	<i>12</i>
3.1.3. <i>Le dispositif d'intégration des parties prenantes à l'évaluation</i>	<i>12</i>
3.2. Pré-évaluation	14
3.3. Critères et questions.....	17
3.4. Designs de recherche.....	20
3.4.1. <i>Complexité et design basé sur la théorie.....</i>	<i>21</i>
3.4.2. <i>Intégration du genre et designs mixtes et participatifs</i>	<i>22</i>
3.5. Techniques de collecte de données et Indicateurs.....	23
3.6. Rapportage, dissémination et utilisation.....	25
4. Comment l'approche de l'évaluation sensible au genre a essaimé sur le terrain du développement durable ?	27
4.1. Les travaux précurseurs.....	28
4.2. Évaluation systémique inclusive pour l'égalité de genre, l'environnement et les voix marginalisées.....	28
4.2.1. <i>Préparation et définition des frontières de l'intervention et de son évaluation.....</i>	<i>30</i>
4.2.2. <i>Déroulement et flexibilité du design et de la collecte de données.....</i>	<i>32</i>
4.2.3. <i>Déroulement et triangulation systémique.....</i>	<i>32</i>
4.2.4. <i>Utilisation et changement culturel</i>	<i>33</i>
5. Conclusion.....	35
Bibliographie	36

Remerciements

Les auteures tiennent à remercier Muriel Janssens (IWEPS) pour son soutien administratif et Évelyne Istace pour son travail d'édition. Nous remercions également Sébastien Brunet, Administrateur général de l'IWEPS, pour ses remarques qui ont nourri ce *Working Paper* ainsi que pour ses encouragements à la poursuite de ce travail au sein des évaluations menées à l'IWEPS.

Nous souhaitons vivement recueillir les commentaires et avis de nos lecteurs afin d'améliorer ce document et de développer les échanges sur cette thématique.

1. Introduction

L'évaluation, par définition normative, consiste à porter un jugement sur la valeur d'une intervention publique sur la base de normes, de critères de jugement spécifiés. En général, la valeur de l'intervention publique est essentiellement déterminée par rapport à la norme de l'atteinte des objectifs qui lui sont assignés.

Cet aspect normatif de l'évaluation est d'autant plus important que l'on s'intéresse aux effets genrés des politiques publiques. En effet, dans nos sociétés, les inégalités de genre (cf. Encadré 1) persistent dans de très nombreux domaines (emploi, salaires, pauvreté, répartition du temps, participation à la prise de décision, etc.), malgré l'existence de tout un arsenal législatif visant à les combattre. Face à de telles injustices, au-delà de l'intégration d'un critère de jugement relatif aux inégalités, le choix du paradigme transformatif, parmi les différents paradigmes qui co-existent en évaluation, semble le plus approprié. Il guide la réflexion sur l'ensemble du processus évaluatif afin que ce dernier intègre la volonté de changement vers une société plus équitable, plus démocratique, plus inclusive.

Encadré 1 : Sexe, genre et intersectionnalité

Sexe et genre, deux concepts qui continuent à être souvent confondus. Le sexe est l'appartenance biologique, un sexe est attribué à toute personne à sa naissance en fonction des organes visibles. En revanche, ce qui donne une réalité au sexe, ce sont les normes de genre (cf. Butler, 1990). « *Le sexe c'est "ce que l'on a", le genre c'est "ce que l'on nous dit d'en faire" mais aussi "ce que l'on en fait"* » (Amsellem-Mainguy, Gelly et Vuattoux 2017). Le genre est donc, pour en donner une définition simple, une construction sociale des rapports entre les sexes et des rôles attachés aux sexes masculins et féminins dans la société.

L'intersectionnalité nous mène un pas plus loin, elle renvoie à l'idée que les différents rapports de pouvoir (en fonction du sexe, des caractéristiques physiques, de l'origine ethnique, de l'âge, de la religion, etc.) sont étroitement imbriqués, que certaines personnes ou groupes subissent simultanément plusieurs formes de domination et de discrimination. Cette notion a été forgée en 1989 par Kimberley W. Crenshaw, afro-féministe américaine, qui a consacré ses recherches universitaires au *Black Feminism*. Ce mouvement féministe américain revendique un point de vue particulier des femmes afro-américaines à la fois sur le féminisme en général et sur les luttes contre la ségrégation entre Blancs et Noirs. Le sens du terme a depuis été élargi, dans les années 2010, avec la montée du cybermilitantisme, et englobe désormais toutes les formes de discriminations qui peuvent s'entrecroiser.

Dans un contexte de besoins illimités et de moyens publics pour y répondre, qui sont par définition limités, il importe d'allouer ces moyens de la façon la plus efficace et efficiente possible, évitant tout gaspillage. D'où l'importance de bien identifier et analyser les besoins auxquels l'intervention publique prétend vouloir apporter une réponse. Les approches *mainstream* qui ignorent les différences existantes entre femmes et hommes conduisent à des politiques mal adaptées ou calibrées et sous-efficaces.

L'intégration d'une perspective de genre est d'ailleurs une obligation depuis 2007 quand la Belgique a adopté sa loi sur le *gender mainstreaming* (approche intégrée de la dimension de genre). Cette stratégie a pour ambition de renforcer l'égalité des femmes et des hommes dans la société, en intégrant la dimension de genre à toutes les étapes de la politique publique : l'élaboration, la mise en place et l'évaluation (Fallon 2004). Une politique qui intègre la dimension de genre est une

politique où la situation des femmes et des hommes concernés a été examinée de manière comparative, où les éventuelles inégalités entre les sexes ont été identifiées et qui cherche à les réduire ou à les éliminer. Le *gender mainstreaming* est ainsi une approche structurelle qui porte sur toutes les phases du « cycle politique », dont celle de l'évaluation, et qui s'impose à tous les acteurs impliqués. Enfin, le *gender mainstreaming* est une approche transversale, c'est-à-dire qu'il s'applique à tous les domaines politiques, « *il exige que les murs verticaux entre les secteurs politiques soient brisés en faveur d'une gouvernance transversale* » (traduit de l'anglais, Woodward 2003). En termes de gouvernance, à l'heure des changements climatiques, de la surexploitation des ressources naturelles et des inégalités entre pays et individus, l'analyse des politiques publiques et leur évaluation se doivent d'intégrer l'ensemble des effets sociaux, économiques, environnementaux qu'elles génèrent, c'est-à-dire opter pour des évaluations sensibles à la durabilité (cf. Encadré 2).

Encadré 2 : Durabilité et développement durable

Le concept de durabilité s'inscrit dans la ligne directe du développement durable (*sustainable development*). La définition la plus communément admise est celle d'un « *développement qui répond aux besoins des générations actuelles sans compromettre la capacité des générations futures de répondre aux leurs* » (Commission mondiale sur l'environnement et le développement, 1988). La popularité du terme développement durable est sans doute liée à la multiplicité des vues autour du terme ou en tous cas de son opérationnalisation.

En langue française, le terme de durabilité peut engendrer une incompréhension car le caractère durable peut s'entendre comme lié à la durée, à une certaine stabilité. Cette interprétation est trop restrictive pour notre propos.

Dans ce chapitre, la durabilité se définit comme le fait de considérer les dimensions sociales, économiques et environnementales (le genre, l'inclusion sociale, le développement économique, les ressources naturelles, etc.) d'une intervention en tenant compte des interdépendances entre ces dimensions et de la temporalité des effets de l'intervention sur celles-ci. La question des éventuels arbitrages entre effets de plus ou moins long terme et entre dimensions est centrale pour l'évaluation sensible à la durabilité.

Comment intégrer ces dimensions transversales dans l'évaluation des politiques publiques ? C'est l'objet de ce *Working Paper* qui s'adresse ainsi à la fois aux commanditaires et aux évaluateurs pour les soutenir dans cette démarche. De nombreux travaux traitent, et depuis longtemps, de la question du genre et de son intégration dans la discipline évaluative. Nous adopterons dès lors en premier lieu des lunettes « genre ». Dans un deuxième temps, nous verrons comment les verres peuvent être adaptés pour adopter cette perspective multi-dimensionnelle et également intersectionnelle de la durabilité.

2. Qu'entend-on par évaluation sensible au genre ?

L'évaluation sensible au genre est une approche évaluative¹ où la focale est mise sur la question du genre. Cette approche en évaluation n'est pas du tout limitée aux interventions qui répondent explicitement à un besoin ou un problème lié au genre, mais peut au contraire s'appliquer aux interventions de tout domaine. Cette intégration du genre dans l'évaluation peut être plus ou moins prononcée et il est intéressant de l'envisager, non de façon binaire mais sous forme d'un continuum².

Figure 1 : Continuum de l'intégration du genre en évaluation



Si on exclut les évaluations aveugles à la question du genre, dans ce continuum, trois catégories d'évaluations peuvent être distinguées pour clarifier les différentes manières d'intégrer le genre :

- **Adaptation** : L'évaluation apprécie l'intervention en prenant en compte les différences et inégalités entre les sexes. Cette manière de faire est relativement économique. Cette identification peut inciter les acteurs à analyser les sources et les conséquences de ces différences et enclencher une démarche vers davantage de sensibilité au genre. Cette adaptation est cependant souvent limitée à la distinction entre femmes et hommes au niveau de la participation à l'intervention et des résultats. Les deux groupes sont visibilisés mais sans analyse de genre et sans qu'ils ne participent au processus évaluatif.
- **Sensibilité** : L'évaluation apprécie l'intervention en analysant les droits, les rôles, les ressources, les responsabilités et les comportements qui sont associés au fait d'être de sexe féminin ou masculin, ainsi que les relations de pouvoir entre hommes et femmes.
- **Transformation** : L'évaluation revendique un agenda politique et normatif de modifications des relations de genre. Les valeurs d'égalité, d'équité, de justice sociale pour les différents sexes et plus largement pour toutes les personnes vulnérables, sont explicites et mobilisées au sein de l'évaluation. D'une part, le contenu de la recherche évaluative va inclure l'analyse des croyances et des mécanismes qui maintiennent ces rapports de pouvoir déséquilibrés

¹ Le terme « approche évaluative » renvoie à la manière dont les évaluateurs et évaluatrices interagissent avec les parties prenantes à l'évaluation (Ridde et Dagenais, 2009, p. 27). Le terme « approche » est aussi utilisé pour mettre en exergue un angle de vue ou un objectif de l'évaluation ou de son processus, comme celui du genre, de la prévention de conflits ou de l'utilisation de l'évaluation (Befani, 2016, p.14).

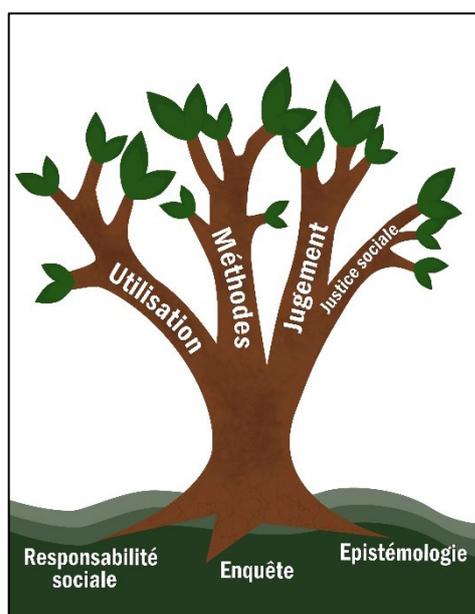
² Adapté : (1) du continuum de genre de l'IGWG (Groupe de travail interagences sur le genre) <https://www.igwg.org/about-igwg/#continuum> (conçu pour comprendre le genre au niveau des programmes, il est ici mobilisé pour l'instrument d'action publique qu'est l'évaluation des politiques publiques) ; (2) de l'échelle d'évaluation genrée (Bamberger et Mabry, 2020, p. 364).

et, d'autre part, le processus évaluatif va avoir pour objectif d'encapaciter les personnes vulnérables ou marginalisées via leur participation dans le processus évaluatif.

Dans le reste de ce chapitre, l'évaluation sensible au genre sera entendue comme toute évaluation qui va au-delà de l'adaptation et qui adopte *a minima* une sensibilité au genre. L'approche suppose que pour réduire les inégalités entre les femmes et les hommes, il est essentiel de prendre en compte les relations de pouvoir existantes. Elle se base sur des contributions féministes au domaine de l'évaluation des politiques publiques, leur focalisation sur la nature systémique et structurelle des inégalités entre les sexes et leur demande pour plus de justice sociale (Bustelo, 2011 ; Sielbeck-Bowen *et al.*, 2002). L'évaluation sensible au genre est un outil clé pour explorer les causes structurelles des inégalités entre les sexes et pour déterminer les implications différentielles des politiques publiques pour les femmes et les hommes (De Waal, 2006).

Cette approche sensible au genre peut trouver des points d'ancrage dans les différents modèles et approches évaluatifs et s'accommoder des différents paradigmes. En se référant à l'« Arbre de l'évaluation »³ (cf. Figure 2), une évaluation sensible au genre dans la branche des méthodes pourrait s'appuyer sur un design basé sur la théorie qui mettrait en exergue les valeurs et les intérêts des publics sous-représentés. Une évaluation sensible au genre dans la branche de l'utilisation pourrait adopter une approche axée sur l'utilisation qui favorise l'implication des personnes dans le processus et offre ainsi l'opportunité du renforcement des capacités de ces personnes. L'approche d'évaluation sensible au genre a toutefois le plus d'affinités avec la branche du jugement.

Figure 2 : Arbre de l'évaluation



Source : Adapté de Alkin (2012) et Mertens (2018)

Au sein de celle-ci, certains théoriciens de l'évaluation, tels Ernest R. House (1980, 1990, 2015), Donna M. Mertens (1999, 2010) et Jennifer Greene (1997), mettent un focus sur la représentation des publics marginalisés et la question de la justice sociale. Cette dernière est le principe prioritaire qui va guider le travail de l'évaluateur. C'est pourquoi Mertens (2018) identifie non trois mais quatre

³ Alkin (2012) et Christie et Alkin (2013) ont utilisé la métaphore de l'arbre pour classer les différents modèles et approches d'évaluation. L'évaluation puise ses racines dans la nécessité de responsabilité sociale, dans l'utilisation des méthodes de recherche en sciences sociales et dans les discussions épistémologiques. Les différents modèles et approches se sont développés en mettant une focale soit sur l'utilisation de l'évaluation (branche de l'utilisation), soit sur les méthodes à mobiliser (branche des méthodes), soit sur le jugement évaluatif (branche du jugement).

branches dans la pratique évaluative : la branche des méthodes, la branche de l'utilisation, la branche du jugement et, enfin, la branche de la justice sociale. Ces branches relèvent respectivement des paradigmes (post-)positiviste, pragmatique, constructiviste et transformatif. Les évaluations adoptant la plus grande sensibilité au genre relèvent de ce dernier paradigme dont les principaux fondements philosophiques sont repris dans l'encadré 3.

Encadré 3 : Le paradigme transformatif et la branche de la justice sociale (Mertens, 2010, 2016, 2018)

Chaque paradigme peut se définir sur la base d'hypothèses ayant trait à quatre domaines : l'ontologie, l'axiologie, l'épistémologie et la méthodologie. Le paradigme transformatif s'appuie sur les hypothèses suivantes :

- **L'hypothèse ontologique sur la nature de la réalité**

Cette hypothèse reconnaît la nature multidimensionnelle de la réalité. Il existe autant de perceptions différentes de la réalité qu'il y a d'individus. L'évaluateur est responsable de rendre visibles ces différentes perceptions de la réalité et les facteurs qui les influencent (genre, caractéristiques physiques, origine ethnique, situation socio-économique, éducation, religion, etc.) afin d'examiner de manière critique les conséquences d'accepter une version de la réalité plutôt qu'une autre. Ceci implique d'encourager les parties prenantes à réexaminer de manière critique leurs propres représentations et hypothèses concernant l'intervention publique et les groupes qu'elle vise à atteindre et de récolter des données directement auprès de ces derniers.

- **L'hypothèse axiologique sur la nature des valeurs et l'éthique**

Cette hypothèse reflète une prise de conscience de l'omniprésence de la discrimination et la responsabilité de l'évaluateur de comprendre les dimensions critiques de la diversité afin de remettre en question les processus sociétaux qui perpétuent un statu quo oppressif et ainsi contribuer à l'identification d'actions permettant d'instaurer une plus grande justice sociale dans la société.

- **L'hypothèse épistémologique sur la nature des savoirs et la relation entre le connaissant et ce qui serait connu (c'est-à-dire entre l'évaluateur et les groupes cibles)**

Le savoir n'est ni absolu ni relatif, il est créé dans un contexte de pouvoir et de privilège. L'évaluateur doit développer des relations de collaboration culturellement sensibles et réactives aux besoins des différentes parties prenantes afin de créer les conditions propices à la révélation de connaissances de différents angles de vue et positions. Des stratégies de communication efficaces sont indispensables pour atténuer les tensions susceptibles de survenir en raison des différences de pouvoir et la nécessité d'aborder des questions sensibles liées à la discrimination et l'oppression.

- **L'hypothèse méthodologique sur la nature de l'enquête systématique**

Aucun choix méthodologique particulier n'est dicté. L'idée est plutôt de privilégier la méthode qui facilitera le mieux (1) l'utilisation du processus et des résultats en vue d'améliorer la justice sociale ; (2) l'identification des forces systémiques qui soutiennent le statu quo et celles qui permettront au changement de se produire ; et (3) la reconnaissance de la nécessité d'une relation critique et réflexive entre l'évaluateur et les parties prenantes.

Le terme de « Justice sociale » est cependant trop limitatif pour couvrir l'ensemble des problèmes épineux auxquels la société fait face. Contribuer à une transformation de la société nécessite la prise en compte de la justice environnementale et économique (Mertens 2016, p. 157-158).

Les freins principaux à une intégration plus systématique et généralisée du genre sont sans aucun doute :

- Le manque de volonté politique : les décisions politiques concernant ce qu'il faut analyser et évaluer ont tendance à marginaliser ce sujet en limitant l'intégration d'une approche transversale genre aux seules évaluations d'interventions qui visent spécifiquement à des transformations de genre ;
- La capacité institutionnelle limitée : l'absence de répartition claire des responsabilités et l'insuffisance des ressources humaines et économiques dédiées aux évaluations sensibles au genre.

Deux éléments permettent cependant d'envisager un futur plus prometteur et font l'objet des deux prochaines sections. Le premier concerne les connaissances accumulées, les outils développés qui sont maintenant à la disposition de la communauté évaluative pour appuyer l'exercice d'évaluations sensibles au genre. Le second a trait au large support politique réservé aux objectifs de développement durable qui incluent un objectif spécifique relatif au genre « Parvenir à l'égalité des sexes et autonomiser toutes les femmes et les filles » tout en envisageant la problématique de façon transversale en combinaison avec d'autres objectifs. Les orientations émises pour évaluer ces objectifs de développement durable reflètent ainsi une prise en compte du genre.

3. Comment faire (rendre) une évaluation « sensible au genre » ?

Intégrer le genre dans l'évaluation nécessite d'adopter des lunettes spécifiques sur l'ensemble du travail évaluatif. Concrètement, cela nécessite :

- de connaître les principales dimensions d'inégalité de genre : la participation des femmes et des hommes aux sphères privée et publique, le contrôle de leur corps par les femmes, les besoins de genre pratiques et stratégiques⁴, les différences au niveau de l'emploi du temps des femmes et des hommes et l'inégal accès et contrôle aux/des ressources et services ;
- d'être conscient du lien fort qui existe entre inégalités de genre et autres formes d'inégalités dans une perspective intersectionnelle (les oppressions ont tendance à se combiner, il y a souvent une confluence de motifs en cause, les caractéristiques physiques, l'origine ethnique, le handicap, etc. venant renforcer les inégalités de genre) ;
- de réviser l'ensemble du processus évaluatif, de sa préparation jusqu'à l'utilisation de l'évaluation, pour rendre l'évaluation la plus sensible possible à la dimension de genre.

La figure ci-après, inspirée par Espinosa (2013), reprend les points d'ancrage essentiels d'une perspective genre dans le processus évaluatif. La suite de la section développe tour à tour ces fondamentaux pour une évaluation sensible au genre.

Figure 3 : Rendre l'ensemble du processus évaluatif « sensible au genre »



⁴ Les besoins pratiques de genre se réfèrent aux besoins qui surviennent en raison du rôle des femmes dans la société mais ne remettent pas en cause ce rôle, tels que les insuffisances dans les conditions de vie, les soins de santé, la situation sur le marché du travail, etc.

Les besoins stratégiques en matière de genre font référence aux besoins que les femmes identifient pour surmonter leur position subalterne dans la société ; elles sont liées à la division sexuelle du travail, du pouvoir et du contrôle, et peuvent inclure des questions telles que les droits légaux, la violence sexiste, l'égalité des salaires, etc. (cf. ILO, 2020, Guidance note 3.1: Integrating gender equality in monitoring and evaluation, v.3, https://www.ilo.org/wcmsp5/groups/public/---ed_mas/---eval/documents/publication/wcms_746716.pdf).

3.1. CADRE DE L'ÉVALUATION

En tant que commanditaire d'une évaluation sensible au genre, il est essentiel de mentionner explicitement ce caractère « sensible au genre » dès l'entame de la démarche évaluative. Certains points doivent être tranchés au niveau de ce qui est attendu de l'évaluation et de la communication de ces attentes aux sous-traitants potentiels (Espinosa, 2013). Le mandat d'évaluation et le cahier spécial des charges méritent d'être réfléchis en fonction d'une série de dimensions différentes qui, toutes, sont importantes pour garantir la sensibilité au genre de l'évaluation qui sera réalisée.

3.1.1. L'équipe évaluatrice

Il importe de préciser quel type de connaissance et d'expérience est attendu (connaissance générale des inégalités de genre, expérience dans le domaine d'analyses de genre, etc.) et s'il s'agit là d'un critère d'exclusion ou d'un atout. Dans certains cas, il peut être envisageable d'organiser une montée en puissance après l'attribution du marché via la participation à des activités de sensibilisation et de formation ou le recours à une expertise externe pointue à des moments définis du processus d'évaluation. Le cahier spécial des charges se doit d'être précis sur ces questions.

Le sexe de l'évaluateur ou de l'évaluatrice externe ou l'équilibre des sexes au sein de l'équipe d'évaluation peut également être un élément pertinent, en particulier lorsque l'évaluation porte sur un sujet sensible (par exemple les violences conjugales). S'il est souhaité que l'évaluation se réalise par une personne d'un sexe donné ou par une équipe caractérisée par une certaine représentation des sexes, le cahier spécial des charges doit veiller à ce que cette attente soit formulée dans le respect du droit du travail.

3.1.2. Le contenu thématique de l'évaluation

Si les évaluations d'interventions visant des objectifs d'égalité de genre intègrent la dimension genre, l'évaluation de toute autre intervention n'aborde pas de façon systématique la question du genre alors que l'intervention pourrait soutenir un *statu quo* oppressif ou produire des effets différenciés selon le sexe. L'intégration de critères d'évaluation genrés et/ou de questions évaluatives genrées permet d'explicitier et de concrétiser la demande d'évaluation sensible au genre.

Si une analyse de l'intervention sous la dimension genre n'a pas encore été réalisée au moment d'établir le mandat et le cahier des charges de l'évaluation, ces documents doivent explicitement demander que la pré-évaluation inclue cette analyse afin de soutenir l'identification des questions évaluatives qui formeront l'ossature de l'évaluation. Murthy (2016) montre que les formes classiques de logique d'intervention, telles que le modèle logique ou la théorie du changement, ont été adaptées au sein de différentes instances pour incorporer la dimension genre. D'autres modèles ont également émergé à l'image de la matrice des changements (cf. *infra*). La logique d'intervention, quelle que soit sa forme, permet de mettre en lumière ou de porter à discussion auprès des parties prenantes des hypothèses quant aux effets de genre possibles, ce qui alimente l'élaboration de questions évaluatives sensibles au genre. Selon le degré d'intégration du genre (cf. Figure 1), les questions pourront porter sur l'intégration du genre dans la conception de l'intervention que ce soit à des fins d'efficacité de l'intervention ou pour répondre aux besoins spécifiques des hommes et des femmes, jusqu'à s'interroger sur la contribution de l'intervention au rééquilibrage des relations de pouvoir basées sur le genre (Chigateri et Saha, 2016).

3.1.3. Le dispositif d'intégration des parties prenantes à l'évaluation

Un principe majeur de l'évaluation sensible au genre est de placer les femmes (en général, le groupe vulnérable en termes de relations de pouvoir fondées sur le genre) et leurs réalités maté-

rielles au centre de l'évaluation. Pour ce faire, l'évaluation sensible au genre s'appuie sur « **un processus inclusif, participatif et respectueux de toutes les parties prenantes** » (ONU Femmes, 2015, p. 4). L'encadré 3 détaille les nombreux avantages de l'intégration des parties prenantes. La participation n'est toutefois pas dichotomique et plusieurs choix doivent être posés et méritent d'être explicités dès le cahier des charges : l'identité des parties prenantes à intégrer, les étapes auxquelles elles vont participer et avec quels rôles/pouvoirs⁵.

Lors de la détermination des objectifs de l'évaluation, une implication des parties prenantes a peut-être déjà eu lieu, auquel cas l'équipe évaluative a tout intérêt à en être informée afin de maintenir les contacts. Dans tous les cas, cette implication est à mettre en place. Une fois l'identification des parties prenantes réalisée, la **sélection d'un nombre restreint de participants à l'évaluation** est nécessaire afin de garantir l'efficacité du dispositif participatif. Le commanditaire et/ou l'équipe évaluative devra peser le pour et le contre entre les avantages et les risques, tels qu'une moindre impartialité, et les ressources additionnelles nécessaires en termes de coût et de temps. L'engagement de ce groupe peut se produire tout au long du processus d'évaluation, à partir de l'identification de ce qui doit être évalué jusqu'à la collecte, l'interprétation et l'utilisation des données et des résultats. L'équipe évaluative doit s'assurer que chacun des membres ait parfaitement compris la manière dont il va contribuer au processus évaluatif. Il est utile de rédiger un document de référence reprenant les **rôles et responsabilités** de chacun au sein du groupe.

Encadré 4 : Les avantages de la participation des parties prenantes dans une évaluation sensible au genre

- **Les contributions** : Les différents points de vue enrichiront la conception et l'approche de l'évaluation et aideront notamment à : identifier des questions liées à l'égalité des sexes et aux droits humains et des implications qui peuvent ne pas avoir été clairement relevées lors de la phase de conception et de mise en œuvre du programme ; identifier et trouver des solutions de manière collective aux contraintes et aux défis de l'évaluation ; faciliter et guider la réalisation de l'évaluation ; et apporter des contributions réalistes et pratiques à l'examen et à l'analyse des résultats, conclusions et recommandations de l'évaluation.
- **Utilisation et utilité** : C'est la mesure où une évaluation est exploitée et utile qui détermine en fin de compte le succès d'une évaluation. Les approches participatives d'évaluation ont démontré la crédibilité grandissante des résultats de l'évaluation et par conséquent son utilisation croissante. L'engagement des parties prenantes dès le début du processus, notamment pour identifier les raisons qui motivent l'évaluation, permet de gérer en amont les attentes sur ce que l'évaluation pourra fournir. La participation des intervenants est fondamentale à l'appropriation et la promotion de l'utilisation de l'évaluation.
- **Renforcement des capacités** : Être impliqué dans le processus d'évaluation est en soi une expérience d'apprentissage et peut servir à renforcer la capacité des parties prenantes à travers une exposition accrue aux questions portant sur l'égalité des sexes et les droits humains et aux approches d'évaluation qui tiennent compte du genre. Ceci peut aider à établir des liens clairs entre les programmes et les grands objectifs de changement sociaux ; encourager les bonnes pratiques dans le suivi et la mesure des progrès réalisés dans le domaine de l'égalité des sexes et des droits humains ; et peut être une étape pour pousser les détenteurs d'obligations (à savoir, les fonctionnaires ou les autorités gouvernementales) à respecter les engagements en matière d'égalité des sexes et des droits humains.

⁵ Pour plus d'informations, voir entre autres les travaux de Monnier (1992) et de Baron et Monnier (2003) sur les approches pluralistes et participatives en évaluation de politiques publiques.

- **Responsabilisation** : Réunir les détenteurs d'obligations et les titulaires de droits génère un cadre de responsabilité mutuelle, de transparence et d'application des principes fondamentaux qui sous-tendent l'égalité des sexes et les droits de la personne humaine.
- **Autonomisation** : L'implication des intervenants et des différents bénéficiaires à toutes les étapes du processus d'évaluation et la mise à leur disposition d'un espace leur permettant de déterminer la manière dont un véritable processus de réflexion et d'évaluation doit être entrepris peuvent leur octroyer les moyens de s'approprier les interventions.

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

Il va de soi que les moyens budgétaires et temporels doivent être en adéquation avec les attentes exprimées dans le cahier spécial des charges.

3.2. PRÉ-ÉVALUATION

Cette étape d'analyse de l'intervention est essentielle pour toute évaluation. Elle permet d'appréhender le contexte dans lequel un problème public a été mis à l'agenda, le diagnostic qui en a été fait, le processus par lequel différents acteurs ont élaboré l'intervention, et le modèle de causalité par lequel l'intervention est supposée atteindre les objectifs fixés. Dans le cas d'une évaluation sensible au genre, cette étape va nécessiter plus ou moins de ressources selon que l'intervention a, ou n'a pas, été conçue de manière sensible avec un objectif d'égalité de genre. Si l'intervention est sensible au genre, un état des lieux dans une perspective de genre aura alors été réalisé et certains mécanismes favorisant l'égalité de genre ou la transformation des relations de pouvoir auront été identifiés. Si l'intervention a été conçue sans prendre explicitement en compte la problématique du genre, l'exercice d'évaluation doit alors servir à révéler les effets positifs ou négatifs des rapports de genre, que ce soit au niveau de la situation initiale ou à celui des effets de l'intervention elle-même.

La reconstruction de la logique d'intervention est une opportunité pour une évaluation sensible au genre, mais ce n'est pas la panacée (Hivos, 2014), car la logique d'intervention pourrait également être aveugle à la dimension genre, en particulier pour ces interventions dont l'objectif premier n'est pas l'égalité de genre. Au-delà des documents de programmation de l'intervention, l'équipe d'évaluation devra s'appuyer sur des ressources théoriques⁶ d'analyse du genre qui seront des outils précieux pour mettre à jour les hypothèses en termes de genre sur lesquelles s'appuie le modèle causal. Par exemple, l'hypothèse d'un lien causal entre la mise en place de cours de français langue étrangère en soirée et la participation des populations ciblées peut soulever de nouvelles questions si une perspective de genre est adoptée : les modalités d'apprentissage sous forme de cours en présentiel sont-elles les plus adéquates pour une population féminine ou masculine ? Les plages de disponibilité sont-elles les mêmes entre hommes et femmes ? Le statut sur le marché du travail ne doit-il pas être pris en compte dans la définition des modalités d'apprentissage ? etc.

Un cadre d'analyse de genre fournit une structure pour organiser l'information sur les rôles et les relations de genre. Il en existe plusieurs⁷ : le cadre analytique de Harvard (ou cadre des rôles genres), la planification axée sur les personnes, le cadre de Moser (ou cadre des triples rôles ou cadre de Moser-Levy), la matrice d'analyse de genre de Rani Parker (1993), la matrice des changements

⁶ Comme le soulignent Albarello, Aubin et Van Haepelen (2016), lors de la phase de pré-évaluation, le recours à des ressources théoriques, autrement dit l'appui sur des courants de pensée qui ont été construits progressivement et qui ont été validés dans différentes disciplines scientifiques, est nécessaire pour conférer un réel socle à l'évaluation. La théorie remplit un rôle de « tiers-distanciant » permettant de clarifier les questions d'évaluation qui, à leur tour, vont constituer l'ossature de la démarche évaluative. La référence à des ressources théoriques contribue ainsi à renforcer et légitimer les résultats ultérieurs de l'évaluation.

⁷ Pour une description détaillée de ces différents cadres et des clés de lecture pour choisir le plus approprié en fonction du contexte et des hypothèses analysées, voir March *et al.* (1999), Warren (2007) et Ligerio *et al.* (2014).

(développée par Rao et Kelleher en 2005, adaptée ensuite par Batliwala en 2008), le cadre d'analyse des capacités et des vulnérabilités, le cadre d'autonomisation (*empowerment*) des femmes ou cadre de Longwe et enfin l'approche des relations sociales. Ces cadres sont des instruments pratiques, initialement conçus pour aider leurs utilisateurs à intégrer une analyse de genre dans leur recherche et/ou conception de politiques, mais qui sont également un support important au niveau de l'évaluation de politiques pour reconstruire la logique d'intervention.

Les différents cadres d'analyse de genre partagent certains traits communs. Par exemple, tous reconnaissent et soulignent l'importance de considérer non seulement le travail rémunéré mais aussi le travail non rémunéré (tâches ménagères et de soins aux personnes). Ils diffèrent toutefois dans leur portée et leur importance et surtout dans leur compréhension de la nature du pouvoir et de l'inégalité et donc dans leurs hypothèses quant à ce qui doit être analysé et adressé. Certains cadres poursuivent plutôt des objectifs d'efficacité, d'autres des objectifs *d'empowerment* ou d'autonomisation des femmes. Pour citer Kabeer (1994 : 303, traduit de l'anglais), « *il y a un lien intime entre façons de penser et façons d'agir* ». Chaque cadre se base sur un nombre limité de facteurs considérés comme étant les plus importants dans l'ensemble extrêmement vaste de facteurs influençant chaque situation et ne peut ainsi produire qu'un modèle grossier de la réalité beaucoup plus complexe.

Le choix du cadre approprié dépendra de la tâche à accomplir, du contexte et des ressources disponibles, des valeurs et hypothèses des concepteurs et évaluateurs de politiques, etc. En pratique, les cadres peuvent aussi être combinés pour permettre leur meilleure adaptation possible à l'évaluation qui les mobilise.

Dans la suite de cette section, nous avons fait le choix de mettre le focus sur deux de ces cadres, dont la combinaison présente, à notre sens, une base solide pour identifier les impacts différents d'interventions sur les femmes et les hommes et ainsi reconstruire la logique d'intervention d'une politique publique d'une manière sensible au genre. Il s'agit de la matrice d'analyse de genre d'une part (cf. Figure 4) et de la matrice des changements d'autre part (cf. Figure 5).

Dans la première matrice, la matrice d'analyse de genre, l'impact d'une intervention est étudié dans quatre domaines : le travail, le temps, les ressources (en tenant compte de l'accès et du contrôle) et les facteurs socioculturels. Dans le domaine du travail, il s'agit d'identifier des changements au niveau des tâches, du niveau de compétence requis et de la capacité productive. Dans le domaine du temps, ce sont les changements au niveau de la durée des différentes tâches qui sont étudiés. La catégorie des ressources fait référence aux changements d'accès aux ressources en conséquence de l'intervention, ainsi que dans l'étendue du contrôle sur l'évolution des ressources de chaque groupe analysé. Enfin, dans le domaine des facteurs socioculturels, il s'agit des changements de rôles ou de statut de genre à la suite de l'intervention, ou encore des changements dans la vie sociale plus largement des groupes cibles.

Figure 4 : Matrice d'analyse de genre

	Travail	Temps	Ressources	Culture
Femmes				
Hommes				
Ménage				
Communauté				

Source : Parker (1993)

La matrice d'analyse de genre peut être utilisée lors de la phase de pré-évaluation comme support à l'élaboration de la logique d'intervention et ce, quel que soit le futur design de recherche de l'évaluation. Elle est souvent utilisée au-delà de l'étape de pré-évaluation comme support à une évaluation participative basée sur la théorie, que ce soit pour évaluer l'impact potentiel, ou par la suite l'impact réel, de l'intervention sur les relations de genre (cf. *infra*).

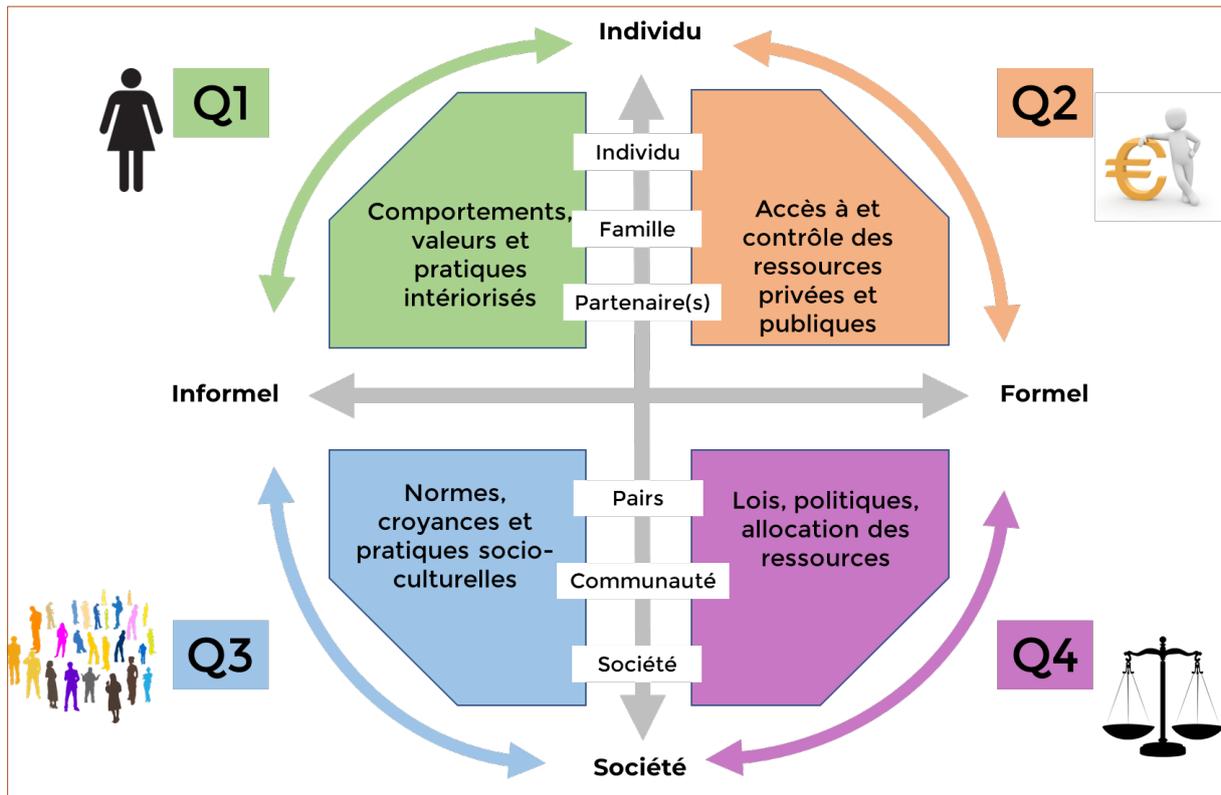
Elle s'appuie sur une technique d'identification et d'analyse des différences entre les sexes qui s'inscrit dans une logique participative. La matrice est remplie (et régulièrement mise à jour afin de compléter progressivement les effets attendus et inattendus) par un groupe au sein de la communauté ciblée par l'intervention qui, de préférence, inclut autant de femmes que d'hommes. Ce cadre s'adapte assez aisément au manque de données quantitatives sur les rôles de genre.

En tant qu'outil d'une évaluation transformative (cf. Figure 1 : Continuum de l'intégration du genre), son utilisation est destinée à amorcer un processus d'analyse et d'évaluation par les membres mêmes de la communauté (ciblée par l'intervention à évaluer), qui sont encouragés à identifier et remettre en question de manière constructive leurs représentations sur les rôles de genre. L'expertise technique de personnes extérieures est considérée non pas comme indispensable mais tout au plus comme facilitant l'analyse de genre. La transformation envisagée doit émaner et être portée par les parties prenantes les plus directement concernées.

La deuxième matrice, celle des changements⁸, est citée dans des travaux universitaires et utilisée par un large éventail d'organisations de toute taille ainsi que par des femmes et des groupes ou communautés locaux dans le monde entier. Elle a été conçue par l'organisation *Gender at Work*, un réseau international de connaissances féministes qui œuvre pour mettre fin à la discrimination à l'égard des femmes et construire des cultures d'inclusion. Les deux quadrants supérieurs se réfèrent à l'individu. À droite, se trouvent des changements dans les conditions individuelles observables, par exemple, l'augmentation des ressources, l'accès à la santé et à l'éducation, etc. À gauche, on retrouve la conscience et les capacités individuelles – connaissances, compétences, conscience politique et engagement à évoluer vers l'égalité. La partie inférieure relève du niveau de la société, du système. Le quadrant de droite fait référence aux règles formelles énoncées dans les constitutions, aux lois et aux politiques. Le quadrant de gauche porte sur l'ensemble des normes discriminatoires informelles et des structures qui maintiennent l'inégalité.

⁸ Initialement conçue par Rao et Kelleher (2005).

Figure 5 : Matrice des changements



Source : Rao et Kelleher (2005, adaptée par Batliwala (2008)

En rendant visibles les différentes dimensions de l'(in)égalité des sexes ainsi que les changements dans les relations de pouvoir entre les sexes sur deux axes variant de l'individu à la société et de la sphère formelle à la sphère informelle, cette matrice permet d'ajouter des niveaux d'analyse comparée à la matrice d'analyse de genre qui, elle, aborde les changements dans les relations de pouvoir dans quatre domaines clés de l'inégalité de genre traversant nos sociétés.

Il nous semble qu'en combinant ces deux cadres d'analyse, l'inégalité de genre et les relations de pouvoir sont analysées dans toutes leurs facettes. En effet, au sein de chaque quadrant de la matrice des changements, elles sont étudiées dans les quatre domaines clés que sont le travail, le temps, les ressources et les facteurs socioculturels.

En tant qu'instruments pratiques, les cadres comme ceux présentés en détail ci-dessus permettent de développer certaines parties de la logique d'intervention, que ce soit au niveau de l'atteinte des publics cibles ou des divers résultats escomptés ou inattendus. Comme dans toute évaluation, les ressources étant limitées, des choix devront être opérés quant aux aspects de l'intervention pour lesquels les besoins d'informations sont essentiels et/ou ceux qui ont le plus de valeur pour le commanditaire et les parties prenantes. Ces choix se refléteront dans les critères d'évaluation et les questions évaluatives.

3.3. CRITÈRES ET QUESTIONS

Une évaluation sensible au genre nécessite une adaptation des critères d'évaluation pour prendre en compte différentes dimensions du genre comme la division genrée des rôles et du travail, la participation des femmes et des hommes aux sphères privée et publique, le contrôle du corps des femmes, des besoins de genre pratiques et stratégiques, l'emploi du temps des femmes et des hommes, l'accès aux ressources et leur contrôle par les deux sexes, etc. Comme l'explique Sierra

Leguina (2000 : 97⁹) : « Une évaluation sensible au genre nécessite une réinterprétation des critères pour permettre une analyse moins centrée sur l'action et plus centrée sur les personnes et les processus » (traduit de l'anglais). Plus précisément, les critères doivent prendre en compte les aspects socioculturels, les structures économiques et politiques générant des inégalités entre les sexes.

Il n'y a pas consensus dans la littérature¹⁰ sur la manière de rendre les critères d'évaluation sensibles au genre. Nous pointons ci-après deux voies possibles.

- Espinosa (2013) décline les critères OCDE/DAC pour prendre en considération le genre. La version ci-après a été adaptée pour prendre en compte la révision des critères de 2019 :
 - **Pertinence** : critère qui se réfère à la mesure dans laquelle les objectifs de l'intervention sont ajustés pour tenir compte des différents problèmes et besoins des femmes et des hommes. Ce critère questionne également si la méthodologie adoptée par l'intervention aide les femmes à percevoir les contraintes et limites qui leur sont imposées et à les surmonter ;
 - **Cohérence** : critère qui fait référence à la compatibilité de l'intervention avec les autres interventions, et en particulier avec les normes et conventions concernant l'égalité des sexes auxquelles le pays ou l'institution adhère ;
 - **Efficacité** : critère qui se réfère à la mesure dans laquelle l'intervention a atteint ses objectifs, notamment en termes de bénéfices obtenus par les femmes et les hommes, ou a induit des résultats du point de vue de l'égalité des sexes, et ce, sans référence aux frais engagés pour les obtenir ;
 - **Efficience** : critère qui se réfère à la mesure dans laquelle les résultats en matière d'égalité des sexes sont atteints à un coût raisonnable ; il questionne également si les prestations ont un coût équivalent pour les femmes et les hommes ; et si celui-ci est réparti équitablement ;
 - **Impact** : critère qui se réfère à la contribution de l'intervention à une politique plus large d'égalité des sexes, aux objectifs sectoriels d'égalité et de progrès vers l'égalité à long terme. La contribution peut être directe ou indirecte et concerner des effets attendus ou inattendus ;
 - **Viabilité/Durabilité**¹¹ : critère qui se réfère au caractère pérenne de la plus-value de l'intervention en matière d'égalité des sexes. L'attention est portée sur les éventuels arbitrages entre les différentes dimensions du développement durable et sur les capacités des systèmes nécessaires à la continuité des bénéfices nets dans le temps. Ceci est lié à l'inclusion des besoins stratégiques de genre dans l'intervention et à l'appropriation de celle-ci par les femmes et les hommes.
- De Waal (2006) propose de combiner les critères d'évaluation classiques et les cinq objectifs du *gender mainstreaming* :
 - **La parité** qui fait référence au nombre de femmes et d'hommes participant ou bénéficiant d'un projet ou d'une action ;

⁹ OCDE/DAC (2019), Better Criteria for Better Evaluation – Revised Evaluation Criteria Definitions and Principles for Use, [revised-evaluation-criteria-dec-2019.pdf \(oecd.org\)](#). Le critère de cohérence a été ajouté mettant ainsi en exergue le besoin de recourir à la pensée systématique et de prendre en considération les dynamiques partenariales et la complexité des interventions. Les définitions de la pertinence et de l'efficacité encouragent une analyse approfondie des problèmes d'équité.

¹⁰ Pour une discussion des différentes propositions de critères d'évaluation afin d'intégrer la dimension de genre, voir Ligerio Lasa *et al.* (2014 : 50-56).

¹¹ L'intégration de l'optique développement durable dans les critères du CAD ne se limite pas à ce critère. L'introduction du critère de cohérence répond également à cet objectif.

- **L'égalité (formelle)** qui analyse si les femmes ont les mêmes opportunités que les hommes ;
- **L'équité (égalité des résultats ou égalité réelle)** qui fait référence à l'équivalence dans l'impact sur la vie des femmes et des hommes, en reconnaissant leurs différents besoins, préférences et intérêts. Ce critère reconnaît que la réalisation de l'égalité des résultats peut exiger différents traitements des femmes et des hommes ;
- **L'autonomisation**, c'est-à-dire le degré de conscience des femmes de leur position de subordination et la mesure dans laquelle elles sont capables de la contrer ;
- **La transformation** qui se réfère à l'inclusion de l'égalité des sexes à l'ordre du jour politique et à l'incorporation de l'approche genre dans toutes les politiques et tous les programmes, ainsi que dans les activités administratives et financières.

Le choix des critères d'évaluation rend explicites les dimensions de l'intervention qui vont être analysées et dont les résultats serviront de base de jugement pour déterminer la valeur de l'intervention. Toutes les évaluations ne s'appuient pas sur la définition de critères d'évaluation mais les questions d'évaluation sont en elles-mêmes porteuses de ces critères.

Les questions d'évaluation, si elles doivent contribuer à construire une évaluation sensible au genre, doivent prendre en compte les inégalités entre les sexes et partir de l'hypothèse que les effets de toute politique publique risquent d'être différents pour les femmes et les hommes. L'équipe évaluative doit parvenir à un consensus parmi les parties prenantes autour des questions à inclure dans l'évaluation. Ces questions peuvent être articulées autour des critères d'évaluation choisis. ONU Femmes (2015) estime qu'en général trois à cinq questions clés liées à chacun des critères choisis permettent d'obtenir une évaluation mieux ciblée.

Le tableau ci-après reprend pour chaque critère OCDE-DAC des questions évaluatives adoptant une perspective de genre.

Tableau 1 : Appliquer les critères d'une manière sensible au genre

Critère	Questions-typiques à se poser pour appliquer le critère d'une manière sensible au genre
Pertinence	<p>L'intervention a-t-elle été conçue de manière à répondre aux besoins et aux priorités de tous les genres ? Si c'est le cas, comment ?</p> <p>Dans quelle mesure le design de l'intervention reflète-t-il les droits des personnes de tous les genres et inclut-il le retour d'information/le point de vue d'un éventail diversifié de parties prenantes locales, y compris les groupes marginalisés ?</p> <p>L'intervention répond-elle aux besoins pratiques et stratégiques de tous les genres ?</p>
Cohérence	<p>Dans quelle mesure le design, la mise en œuvre et les résultats de l'intervention sont-ils cohérents avec les lois internationales et les engagements en faveur de l'égalité de genre et des droits, y compris la Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes (CEDAW), la Déclaration et le Programme d'action de Beijing, le Programme d'action de la Conférence internationale sur la population et le développement et l'Agenda 2030 ?</p> <p>Dans quelle mesure l'intervention soutient-elle la législation et les initiatives nationales visant à améliorer l'égalité de genre et les droits humains ? Quelles leçons peuvent être tirées ?</p>

Efficacité	<p>L'intervention a-t-elle atteint ses objectifs et les résultats escomptés de manière à contribuer à l'égalité de genre ? Si c'est le cas, comment ?</p> <p>Y a-t-il eu des résultats différents pour différentes personnes ? Si oui, comment et pourquoi ? Des approches différentes étaient-elles nécessaires pour atteindre des personnes de genres différents ? Le suivi et l'analyse des effets différentiels étaient-ils suffisants ? L'intervention a-t-elle été ajustée pour répondre à des préoccupations éventuelles et maximiser l'efficacité ?</p> <p>La théorie du changement et le cadre des résultats ont-ils été alimentés par une analyse en termes d'égalité de genre, d'économie politique et de droits humains ? Si oui, dans quelle mesure ?</p> <p>Dans quelle mesure et pourquoi l'efficacité est-elle différente pour les personnes de genres différents ?</p>
Efficience	<p>Les différentes ressources ont-elles été allouées de manière à tenir compte de l'égalité de genre ? Si oui, comment ont-elles été réparties ? L'allocation différentielle des ressources était-elle pertinente ?</p> <p>Les coûts d'investissement par personne ciblée répondent-ils aux besoins différenciés des personnes de genres différents ?</p>
Impact	<p>Y a-t-il eu des impacts égaux pour les différents genres ou y a-t-il eu des différences liées au genre au niveau de la participation, de l'expérience et des impacts ? Si oui, pourquoi ces impacts différentiels se sont-ils produits ?</p> <p>Dans quelle mesure les impacts liés au genre se recoupent-ils avec d'autres barrières sociales, notamment la race/l'origine ethnique, le handicap, l'âge et l'orientation sexuelle pour contribuer à des expériences et des résultats différents ?</p> <p>Comment les normes et les barrières genrées au sein de l'environnement politique, économique, religieux, législatif et socioculturel plus large ont-elles eu un impact sur les résultats ?</p> <p>Dans quelle mesure les impacts ont-ils contribué à l'égalité des relations de pouvoir entre les personnes de genres différents et au changement des normes et systèmes sociaux ?</p>
Viabilité/ Durabilité	<p>L'intervention a-t-elle contribué à une plus grande égalité de genre au sein des systèmes juridiques, politiques, économiques et sociaux plus larges ? Si oui, comment et dans quelle mesure ? Cela a-t-il entraîné des changements durables de normes sociales qui sont préjudiciables aux personnes de tous ou de certains genres ? Si cela n'a pas pu être atteint, pourquoi pas ?</p> <p>Les réalisations en matière d'égalité de genre persisteront-elles après la fin de l'intervention ? Les processus ont-ils contribué à maintenir ces avantages ? Des mécanismes ont-ils été mis en place pour soutenir la réalisation de l'égalité de genre dans le plus long terme ?</p>

Source : OCDE (2021:33)

3.4. DESIGNS DE RECHERCHE

L'évaluation sensible au genre est un exercice où la sensibilité au genre doit imprégner chacune des étapes de la démarche évaluative, en ce compris l'élaboration du design de la recherche évaluative. Plus qu'un nouveau type de design, il s'agit d'élaborer un design en veillant à l'adéquation tant technique que politique, c'est-à-dire en prenant en compte le contexte évaluatif, les questions

évaluatives, les designs potentiels et la nature de l'intervention. Il n'y a donc pas de design de recherche spécifique aux évaluations sensibles au genre, ni un design-type qui conviendrait à toute évaluation sensible au genre. Toutefois, deux éléments caractéristiques influenceront généralement le design de recherche. D'une part, le caractère (le plus souvent) complexe de l'intervention et de son évaluation, ce qui poussera l'équipe évaluative à envisager une approche évaluative basée sur la théorie. D'autre part, le niveau d'intégration du genre dans l'évaluation car au plus l'évaluation sera sensible au genre, au plus une méthodologie mixte avec un large apport de designs qualitatifs et participatifs seront nécessaires.

3.4.1. *Complexité et design basé sur la théorie*

Il est coutumier de distinguer les interventions simples, compliquées et complexes (Glouberman et Zimmerman, 2002, cité par Rogers, 2008, p. 31). La complexité provient de quatre sources essentielles : l'incertitude, l'émergence, la rétroaction¹² (*feed-back*) et l'intersectionnalité (cf. Encadré 1). Les évaluations sensibles au genre, qu'elles portent sur des interventions qui visent un objectif de genre ou sur toutes autres interventions qui influencent, par leur mise en œuvre ou leurs effets, l'égalité de genre, font face à ces éléments de complexité. Il est essentiel lors de l'analyse de l'intervention d'appréhender les différents domaines et niveaux de complexité pour élaborer une démarche évaluative et un design de recherche pertinents et réalistes. Bamberger, Vaessen et Raimondo (2016) adoptent une vision très large de la complexité qui recouvre tant la complexité de l'intervention vue comme un système que la complexité de l'évaluation. Ils identifient cinq dimensions de complexité¹³, toutes interconnectées. Nous présentons pour chacune d'elles un exemple qui illustre la difficulté accrue en termes d'évaluation :

- La nature de l'intervention
Exemple : cette dimension recouvre par exemple les objectifs de l'intervention qui peuvent être multiples et qui peuvent ou non inclure un objectif d'égalité de genre.
- Les institutions et parties prenantes
Exemple : Certaines parties prenantes à la conception, la mise en œuvre ou l'évaluation de l'intervention peuvent considérer que l'intervention n'a pas d'influence sur l'égalité de genre. Les parties prenantes peuvent également ne pas partager la même vision des mécanismes par lesquels l'intervention va fonctionner. Les parties prenantes peuvent enfin identifier des effets non escomptés.
- La causalité et le changement
Exemple : Le processus de changement attendu d'une intervention en termes d'égalité de genre est souvent long, se modifie dans le temps, connaît des effets de rétroactions, ce qui rend l'évaluation de certaines questions causales très difficile. L'ensemble des effets sont-ils identifiés ?
- L'intégration de la nature du système
Exemple : Une intervention s'inscrit dans un contexte culturel avec ses propres normes et croyances.
- L'évaluation
Exemple : Comment rendre explicites les valeurs de l'équipe d'évaluation et des parties prenantes qui influencent les caractéristiques de l'évaluation ?

Les designs expérimentaux et quasi-expérimentaux, utilisés de façon isolée, sont insuffisants pour saisir les liens complexes entre, d'une part, les ressources et activités, les réalisations et les effets

¹² L'incertitude a trait au manque de contrôle sur des éléments qui affectent l'intervention et les effets de celle-ci. L'émergence concerne ce qui se passe effectivement (que ce soit quelque chose de prévu ou pas) ou ce qui pourrait se passer. La rétroaction peut prendre diverses formes : renforcement, résistance, repoussoir, contre-pied, etc.

¹³ Pour une présentation plus détaillée, voir Bamberger, Vaessen et Raimondo (2016, pp.12-22)

et, d'autre part, le contexte et les structures. Ces designs ne peuvent s'adresser qu'à une forme de complexité restreinte. Les approches évaluatives basées sur la théorie sont dès lors souvent mobilisées pour évaluer des interventions complexes de changement sociétal (Fulbright-Anderson, Kubisch et Connell, 1998) où les effets sont incertains et peuvent émerger à différents niveaux. L'évaluation doit permettre de vérifier si les conditions adéquates pour obtenir l'impact attendu sont bien en place (Kalpazidou Schmidt & Cacace, 2017) et que le processus de production de l'effet attendu opère effectivement. L'évaluation basée sur la théorie considère justement l'intervention comme « une chaîne causale allant de l'intervention initiale à l'impact final, dont le processus n'est pas linéaire, mais instable et non maîtrisé par le décideur public » (Devaux-Spatarakis, 2014, p.54). Les évaluations basées sur la théorie vont donc au-delà de la causalité purement contrefactuelle dans une tentative de reconstruire les mécanismes d'impact de manière détaillée, en impliquant toutes les parties prenantes. Les interventions reposent sur des hypothèses explicites formulées par les parties prenantes qui nous permettent de comprendre ce qui se fait dans la « boîte noire » et pourquoi, ce qui fonctionne ou ne fonctionne pas, pour qui, dans quel contexte et, en cas de non-atteinte des effets attendus, ce qui doit être ajusté (Chen, 1990 ; Pawson et Tilley, 1997 ; Barnes, Matka et Sullivan, 2003). Les évaluations basées sur la théorie cherchent donc à élucider comment et sous quelles conditions une intervention aboutit aux résultats escomptés et non escomptés observés (Döring & Bortz, 2016). Elles explorent non seulement si l'intervention fonctionne, mais aussi comment, pour qui et dans quel contexte (Van Belle, Marchal, Dubourg & Kegels, 2010).

3.4.2. *Intégration du genre et designs mixtes et participatifs*

Plus le niveau de sensibilité au genre de l'évaluation est élevé (cf. Figure 1), plus l'évaluation vise des objectifs larges, allant du constat jusqu'à l'explication, de la compréhension jusqu'à la transformation des relations de genre, en ce compris via le processus évaluatif.

Pour atteindre l'objectif de compréhension, en présence d'interventions complexes, les méthodes quantitatives d'exploitation de données administratives ou d'enquête s'avèrent souvent insuffisantes. En effet, les méthodes mixtes (combinant des données et méthodes quantitatives et qualitatives) sont largement reconnues comme plus appropriées aux évaluations sensibles au genre « pour associer une compréhension des expériences vécues par les femmes et les hommes dans différents types de ménages, communautés et activités économiques (évaluation qualitative) et une estimation de la portée et de la représentativité des observations et la signification statistique des différences observées (évaluation quantitative) » (Bamberger, Segone et Tateossian 2017: 25). Les designs mixtes reposent ainsi sur le recours à une diversité d'outils, tant quantitatifs (enquêtes par questionnaire, données d'enregistrement administratif, économétrie, etc.) que qualitatifs (entretiens individuels, *focus groups*, observation directe, études de cas, etc.).

Pour atteindre l'objectif de transformation des relations de genre, en particulier via le processus évaluatif, un design mixte associant méthodes quantitatives et qualitatives n'est toujours pas en soi suffisant. Encore faut-il que la démarche évaluative intègre une dimension participative. En faisant appel à une large gamme de parties prenantes (public(s) cible(s), décideurs, responsables de la mise en œuvre, ainsi que d'autres acteurs favorables et défavorables à l'intervention), les méthodes participatives permettent d'améliorer la qualité de l'évaluation par la prise en considération d'un large ensemble de points de vue tant sur le choix des questions évaluatives prioritaires que lors de la phase d'analyse des données et ensuite de l'élaboration des recommandations. Les designs participatifs soutiennent le rassemblement de données complexes et riches et la diversité des informations récoltées (en complétant les données administratives et d'enquête avec des données plus subjectives sous forme d'avis, de perceptions, de représentations, etc. de toutes les parties prenantes au processus). La participation des femmes doit s'envisager au plus tôt de la démarche éva-

lative afin d'initier un processus de sensibilisation et d'apprentissage par les échanges d'informations et le questionnement qui s'opèrent en cours d'évaluation. C'est ce processus qui est susceptible d'encapaciter les participantes et de promouvoir des changements à plus ou moins long terme et par différents canaux en faveur de l'équité de genre.

À titre d'exemple, reprenons la combinaison de cadres proposée ci-dessus (cf. Section 3.2), la matrice d'analyse de genre et la matrice des changements. Leur construction intervient en plusieurs phases. Lors de la phase de pré-évaluation, un premier remplissage permet de rassembler l'ensemble des informations disponibles sur l'intervention et son contexte et, éventuellement, selon le degré de mise en œuvre de l'intervention, les premiers éléments de réflexion ou de retour du terrain. Une fois la matrice remplie par les membres du groupe concerné par l'intervention, elle peut être utilisée pour poser la réflexion quant à la désirabilité des effets identifiés, à leur conformité aux objectifs de l'intervention, aux facteurs qui entravent ou soutiennent les effets désirables, aux effets de l'intervention sur les personnes/groupes non directement ciblés et/ou des personnes ciblées plus vulnérables, aux résultats inattendus, etc. Les différentes parties prenantes peuvent en tirer des leçons et des conclusions qu'ils appliqueront dès ce moment dans leur pratique. Ce cadre peut ainsi aider à anticiper de possibles résistances et encourage la prise en compte du soutien à offrir aux personnes à risque. En cours d'évaluation, ces matrices peuvent être remobilisées et actualisées afin d'améliorer, d'une part, la compréhension de l'intervention et, d'autre part, les compétences des parties prenantes à faire entendre leur voix, leurs arguments, à élaborer des propositions, etc. Ces cadres sont en effet flexibles et conçus pour s'adapter aux changements dans le temps, y compris ceux qui sont inattendus. Enfin, lors de la phase finale de l'évaluation, après avoir rempli les cases de la matrice avec les (derniers) changements apportés à l'intervention, les matrices permettent aux membres du groupe de qualifier les résultats/effets identifiés : un signe positif (+) si le résultat est cohérent avec les objectifs, un signe négatif (-) si le résultat est contraire aux objectifs et un point d'interrogation (?) s'ils ne savent pas. Ces signes ne sont pas destinés à être additionnés dans le but de déterminer l'effet net de l'intervention. Cela simplifierait à l'excès l'image d'une réalité complexe et dénaturerait le mélange d'effets positifs et négatifs de toute intervention.

À côté des avantages évoqués, la principale limite tient sans doute au fait qu'en tant qu'outils participatifs, ces matrices permettent certes la prise en compte du point de vue des femmes, mais nécessitent une mise en application prudente qui ne crée pas d'attentes auxquelles le travail évaluatif ne pourra répondre.

3.5. TECHNIQUES DE COLLECTE DE DONNÉES ET INDICATEURS

Comme aux autres étapes, l'intégration du genre impose certains points d'attention au niveau des techniques de collecte d'informations. En particulier :

- Les femmes et les hommes doivent être inclus dans l'échantillon ou la population étudiée et les obstacles potentiels à la participation des femmes doivent être surmontés ;
- La disponibilité de temps différente des femmes et des hommes et leurs divers besoins et intérêts doivent être pris en compte ;
- Une équipe d'évaluation ayant des compétences pour travailler avec des hommes, des femmes et des groupes diversifiés est nécessaire ;
- L'équipe doit avoir les compétences nécessaires pour renforcer la participation et l'appropriation de la pratique d'évaluation et non renforcer le leadership traditionnel et la concentration du pouvoir.

Des « indicateurs sensibles au genre » ou « indicateurs de genre » mesurent spécifiquement comment l'intervention promeut l'égalité des sexes. Ce type d'indicateur permet de suivre comment l'intervention affecte le statut et la position des femmes et des hommes et les relations entre les sexes. Les indicateurs sensibles au genre aident à expliquer les manifestations de l'inégalité entre les sexes qui sont souvent invisibles dans les indicateurs traditionnels. Par exemple, en ce qui concerne l'emploi, les indicateurs de genre incluent tous les types de travail des femmes, rémunéré et non rémunéré, souvent non pris en compte dans les indicateurs traditionnels. L'encadré ci-après développe un autre exemple dans le domaine de la pauvreté.

Encadré 5 : Indicateur relatif au risque de pauvreté

La recherche traditionnelle sur le sujet de la pauvreté table sur un instrument très pratique qu'est « la fonction d'utilité du ménage ». Les mesures de la pauvreté sont basées sur les ressources des ménages dont on suppose qu'elles sont partagées de façon égale entre les différents membres du ménage, de sorte que chacun ait le même statut face à la pauvreté. Le taux de risque de pauvreté standard est calculé comme le pourcentage de personnes vivant dans des ménages dont le revenu disponible équivalent est inférieur à 60% de la médiane de ce revenu dans le pays. Il est donc estimé en faisant l'hypothèse d'une mise en commun et d'un partage de tous les revenus entre les membres d'un ménage, et cela, quel que soit leur apport propre. Selon cette approche, une personne est pauvre si elle appartient à un ménage pauvre, quels que soient ses revenus personnels. Le fait de baser la mesure du taux de risque de pauvreté auquel fait face une personne sur les revenus du ménage auquel elle appartient est problématique, particulièrement pour les femmes. L'hypothèse de mise en commun et de partage des ressources au sein des ménages masque la précarité des femmes et entraîne la mise en place de politiques sociales qui n'adressent pas leur précarité particulière et qui sont donc inefficaces.

La disponibilité de statistiques ventilées par sexe dans tous les domaines, pas uniquement ceux où la dimension de genre est évidente, est un prérequis pour la mise en œuvre d'évaluations sensibles au genre. En l'absence de statistiques différenciées par sexe permettant de mesurer la situation des femmes et des hommes et leurs rôles dans la vie sociale, économique et politique préalablement à l'intervention, l'évaluation des progrès accomplis vers la réalisation de l'égalité des sexes et la hiérarchisation des actions visant à résoudre les questions d'inégalité entre les sexes sont particulièrement difficiles. Un travail de collecte additionnelle de données sera alors nécessaire. En raison de son importance pour faire progresser le statut des femmes, l'élaboration de statistiques différenciées par sexe et d'indicateurs de genre a d'ailleurs été identifiée comme une priorité dans la Plateforme d'action de Pékin de 1995.

En outre, même si des statistiques par sexe existent, il importe de comprendre comment elles ont été construites. Au-delà de l'exemple du taux de risque de pauvreté (cf. Encadré 5), trois problèmes conceptuels et méthodologiques peuvent potentiellement survenir :

- Premièrement, « sexuellement neutre » peut être synonyme de « *gender-blind* ». Lorsqu'une loi ne fait aucune distinction entre les sexes, elle s'applique de manière égale aux femmes et aux hommes. Toutefois, vu que les femmes et les hommes se trouvent dans des situations différentes, une application égale peut désavantager l'un des deux groupes. Le fait de ne pas faire de distinction n'offre donc aucune garantie que le résultat soit neutre, non biaisé. Neutralité n'est pas synonyme d'égalité. Ces biais ne s'appliquent pas qu'à des lois mais aussi à la manière de recueillir des données statistiques. Par exemple, ne pas détailler les données de la même manière pour les hommes et les femmes est potentiellement source de biais.

Un exemple typique est donné par les catégories professionnelles qui, dans les classifications habituelles, tendent à être beaucoup plus détaillées pour les professions dans lesquelles les hommes sont surreprésentés que pour celles où les femmes forment une majorité. Ce même problème s'observe aussi pour les filières d'études et les secteurs d'activité.

- Deuxièmement, des biais peuvent résulter du choix du niveau de mesure des données. Si les données concernant les ménages sont particulièrement intéressantes du point de vue du genre, vu que les inégalités entre femmes et hommes dans différents domaines (emploi, violence, pauvreté, etc.) sont souvent étroitement liées à la situation familiale, elles deviennent un problème si les données ne permettent pas une ventilation par sexe. Vu que les analyses de genre nécessitent des statistiques ventilées par sexe, les données relatives aux ménages constituent un problème. Il n'est par exemple pas possible d'établir, sur la base de l'enquête sur le budget des ménages, qui dépense combien, pour quoi, et quelles sont les différences sexuées au niveau de la consommation. Aussi, dans l'enquête SILC, la possession d'un bien, aussi individuel qu'il soit, tel qu'un GSM par exemple, est une question adressée au ménage et non pas à l'individu : le ménage possède-t-il un GSM ? C'est aberrant et problématique. Et c'est pourquoi il importe de travailler avec deux niveaux de mesure, en complétant les données recueillies au niveau des ménages avec des données individuelles.
- Troisièmement, les enquêtes relatives aux ménages utilisent des personnes de référence ou des proxys, une pratique qui est, elle aussi, potentiellement source de biais de genre. Dans les deux cas, des informations sont récoltées de manière indirecte avec donc un risque réel de déformation :
 - La personne de référence est une personne qui « parle » pour l'ensemble du ménage. Le concept de « personne de référence » est une reformulation plus neutre du point de vue du genre de l'ancien concept de « chef de famille ». Bien qu'il s'agisse d'une adaptation positive, elle ne résout pas tous les problèmes méthodologiques sous-jacents. Pour choisir la personne de référence au sein d'un ménage, plusieurs enquêtes utilisent des critères qui ne sont pas sexuellement neutres, tels que l'âge, le fait d'avoir un emploi ou le plus haut revenu, ce qui aboutit à une surreprésentation de répondants masculins dans l'échantillon utilisé. En revanche, la méthode de l'anniversaire, qui consiste à sélectionner, dans un ménage, la dernière personne qui a fêté son anniversaire, est une méthode neutre du point de vue du genre.
 - Un proxy est un membre du ménage qui répond à la place du répondant absent. Afin d'éviter tout biais, il importe de limiter leur utilisation pour répondre à des questions très concrètes, concernant des faits objectifs, auxquelles il est en principe très facile de répondre. Il faut éviter que des proxys soient utilisés pour répondre à des questions subjectives, qui concernent la satisfaction, la qualité, la perception, les opinions et les sentiments.

3.6. RAPPORTAGE, DISSÉMINATION ET UTILISATION

Le rapport final d'une évaluation sensible au genre va nécessairement refléter la spécificité d'une évaluation axée sur le principe d'égalité des sexes, au niveau des thèmes abordés mais également du dispositif d'intégration des parties prenantes. Ce dernier sera explicité de façon à démontrer sa potentielle capacité à renforcer la position de certains acteurs et la façon dont il évite la survenue de discriminations supplémentaires et/ou l'exacerbation de relations de pouvoirs inégales (ONU Femmes, 2015 : 91).

Le rapport commence par la présentation d'une analyse genrée de la situation problématique ou des besoins à l'origine de l'intervention, d'une part, et de l'intervention évaluée, d'autre part. Il détaille la manière dont le design de recherche a intégré la dimension genre. Il discute explicitement de la capacité de l'intervention à aborder de manière adéquate les questions de genre, de la sensibilité de l'intervention aux questions de genre et formule des conclusions à cet égard. Il se termine par des recommandations sur la façon d'améliorer les performances en matière d'égalité des sexes. Il est par ailleurs opportun que le rapport mette en évidence les enseignements plus généraux et les bonnes pratiques en matière d'égalité des sexes qui sont pertinentes au-delà de la portée immédiate de l'intervention (ILO, 2020).

Une fois le rapport terminé, partagé avec les parties prenantes, la dissémination des résultats et des recommandations est un moment charnière pour encourager l'apprentissage et la transformation. Une attention particulière doit être apportée au suivi des recommandations. Préparer et organiser ce suivi enclenche la réponse des différents responsables aux recommandations de l'évaluation. La conception d'un plan de communication, qui tient compte des modèles et des espaces de communication spécifiques aux femmes et aux hommes (communication plus ou moins formelle, via certaines presses ou médias sociaux, via certains lieux comme les devant d'école ou des clubs sportifs, etc.), est essentielle pour garantir que le rapport atteigne toutes les personnes impliquées et qu'elles puissent toutes s'en emparer et se l'approprier. L'engagement avec les médias, le réseautage contribuent aussi à l'efficacité des efforts visant à (re)façonner la politique (Kelly, 2015, citée dans Mertens, 2018 :104).

La présente section a brossé un portrait synthétique des connaissances accumulées pour incorporer une perspective de genre dans l'évaluation des politiques publiques. La mise en pratique de telles évaluations est cependant encore souvent limitée à des évaluations d'interventions qui visent explicitement un objectif de genre. Les concepteurs et acteurs de mise en œuvre sont en général très réticents à accepter une évaluation – un jugement sur la valeur – de leur intervention qui prendrait en considération la contribution à des objectifs qui ne lui sont pas explicitement assignés. Comment dès lors intégrer le genre comme dimension transversale dans les évaluations. Hunt & Brouwers (2003) mentionnaient déjà l'importance de construire un partenariat entre les institutions internationales, les financeurs d'interventions, les acteurs de l'évaluation et les différentes parties prenantes sensibilisées à l'égalité de genre. À cet égard, l'Union européenne a émis un document de soutien à l'élaboration d'évaluations incluant le genre comme dimension transversale s'appuyant sur le fait que l'égalité des sexes et l'autonomisation des femmes et des filles sont des droits humains fondamentaux et font partie des valeurs fondatrices du Traité sur l'Union européenne et de la Charte des droits fondamentaux (UE, 2018). Le groupe évaluation des Nations unies a quant à lui adopté comme standard de qualité pour les évaluations le fait d'adopter une approche fondée sur les droits de l'homme et une stratégie d'intégration de la dimension de genre (UNEG, 2016, p.24). Ces différentes directives prennent malheureusement place dans un contexte évaluatif focalisé sur des interventions externes d'aide aux pays en transition ou en développement et sont peu mobilisées pour soutenir des évaluations sensibles au genre pour des interventions internes. Le large support politique réservé aux objectifs de développement durable qui incluent un objectif spécifique relatif au genre apparaît comme une opportunité pour transformer les systèmes d'évaluation dans nos pays et les rendre plus sensibles au genre mais également aux enjeux environnementaux. Dans la prochaine section, nous synthétiserons les développements en cours pour dessiner des évaluations sensibles au développement durable. Le lecteur pourra se rendre compte des défis communs entre sensibilité au genre et sensibilité au développement durable et des similitudes des réponses apportées.

4. Comment l'approche de l'évaluation sensible au genre a essaimé sur le terrain du développement durable ?

Si le concept du genre fait l'objet d'une définition largement partagée (cf. Encadré 1), le concept de la durabilité mobilisé en évaluation doit encore être éprouvé. Le rapport Bruntland a posé le cadre contemporain de la durabilité, d'une part, en définissant le développement durable comme « *le développement qui rencontre les besoins actuels sans compromettre la capacité des générations futures à répondre à leurs propres besoins* » (CMED, 1988) et, d'autre part, en appuyant celui-ci sur les trois piliers que sont l'environnement, l'économie et le social. L'évaluation ne découvre pas ce terme de durabilité à la suite de la mise en place du cadre des objectifs de développement durable. Parmi les critères d'évaluation DAC-OCDE, celui de durabilité y figure depuis le départ. Toutefois, son sens a évolué pour s'adapter au contexte actuel : d'une conception statique où il était essentiellement question de la viabilité financière de l'intervention à une conception plus dynamique qui encourage l'analyse « *des arbitrages potentiels entre priorités et de la résilience des capacités/systèmes permettant la continuation dans le temps des bénéfiques* » (OCDE-DAC, 2019). Si la discipline évaluative a déjà touché à ces questions d'arbitrage entre priorités, force est de constater que les évaluations actuelles sont systématiquement réalisées au niveau d'une thématique, d'un secteur et renforcent un fonctionnement en silos. À l'image du genre, la durabilité doit être une dimension transversale, prise en considération par toute évaluation.

Parmi les trois piliers, ceux de l'économie et du social, apparentés au système humain, sont globalement bien traités par les évaluations, même si la dimension de genre ou l'attention aux personnes plus vulnérables sont moins présentes. Concernant le pilier de l'environnement, apparenté au système naturel, à part pour les interventions dont l'objectif prioritaire est environnemental, les évaluations peinent à le prendre en considération : une gestion axée sur les résultats ne favorise pas la prise en considération d'objectifs non assignés explicitement à l'intervention et la connaissance des méthodes en sciences du milieu naturel est peu présente parmi la communauté évaluatrice.

L'Agenda 2030 et les dix-sept Objectifs de développement durable (ODD) ont poussé cette dernière à réfléchir sur ses pratiques et en particulier sur la manière d'intégrer plusieurs dimensions dans l'évaluation. Si le genre, l'environnement et l'inclusion sociale sont des thèmes étudiés dans les évaluations, ce qui est neuf, c'est que les ODD reconnaissent la nécessité d'intégrer des considérations socio-écologiques dans les interventions sociales. En effet, le terme « environnement » est un terme chapeau qui couvre des systèmes humains et naturels couplés, voire imbriqués. Les systèmes écologiques changeants s'entrecoupent ainsi avec les interventions sociales qui visent à (re)façonner le mode de vie des populations, en particulier les plus vulnérables. D'où l'importance de voir l'environnement comme une partie prenante à l'évaluation, sur un pied d'égalité avec les autres, afin de décloisonner et d'ouvrir l'évaluation à des impacts émergents qui autrement seraient restés cachés. Il s'agit ici non pas de révéler, d'analyser et de changer les rapports de pouvoir entre femmes et hommes mais entre l'être humain et les autres formes de vie sur notre terre, la nature, l'environnement, vus comme des sujets méritant reconnaissance et respect plus que comme des objets que l'être humain peut manipuler. Pour reprendre Pelizzon et Gagliano (2015) : « ... *the environment as a marginalized stakeholder, lacking voice and diminished by its non-human status* » [« ... l'environnement en tant que partie prenante marginalisée, sans voix et affaiblie par son statut non humain »].

Quelles sont les bases pour développer une évaluation sensible à la durabilité ? D'une part, des travaux précurseurs existent et connaissent une popularité croissante et, d'autre part, le développement de cadres internationaux pour atteindre des objectifs au niveau planétaire a suscité la réflexion sur leur évaluation. Ceux-ci ont notamment débouché sur une nouvelle approche, l'évaluation systémique inclusive pour l'égalité de genre, l'environnement et les voix marginalisées (Inclusive Systemic Evaluation Fo(u)r Gender, Environnement and Marginalized voices – ISE4GEMs).

4.1. LES TRAVAUX PRÉCURSEURS

Le fondement essentiel d'une évaluation sensible à la durabilité est l'analyse conjointe des deux systèmes : humain et naturel. Des travaux précurseurs, notamment ceux d'E. Ostrom (1990, 2009), sur l'évaluation de la durabilité des systèmes socio-écologiques ou de W. Clark (2007) sur les systèmes de connaissances pour le développement durable, offrent des cadres théoriques soutenant la réalisation de ce type d'évaluation. Rowe (2012) en esquisse les éléments-clés de mise en œuvre :

- Partir de l'hypothèse que vous évaluez une intervention qui s'inscrit dans les systèmes humain et naturel : ce n'est qu'après enquête que cette hypothèse pourrait être levée ;
- Considérer le mécanisme-clé qu'est la connectivité entre les deux systèmes : la logique d'intervention doit s'affranchir des échelles temporelles et spatiales du système humain pour prendre également en considération celles du système naturel ;
- Négocier avec les parties prenantes pour ne pas limiter l'évaluation de l'intervention aux seuls objectifs prioritaires assignés et englober des objectifs de durabilité, même si l'intervention n'a pas été conçue pour ce faire, en ce compris en termes de mobilisation de ressources ;
- S'assurer de disposer d'une équipe évaluatrice qui dispose des connaissances et compétences nécessaires aux deux systèmes : il est important que des experts du système naturel soient impliqués dès la formulation des questions d'évaluation et l'élaboration du design de recherche ;
- Définir le champ de l'évaluation en prenant en compte les différentes échelles géographiques et temporelles des deux systèmes et les différentes unités d'analyse : un arbitrage des échelles et unités d'analyse les plus pertinentes sera nécessaire ;
- Identifier et inclure dans le processus toutes les parties prenantes de l'intervention, en ce compris des représentants/experts du système naturel et veiller à une représentation équilibrée des différents intérêts ;
- Adopter un processus participatif pour veiller à la crédibilité de l'évaluation aux yeux des principaux utilisateurs : s'assurer de leurs questions prioritaires, de la bonne compréhension et acceptation d'un langage ou de méthodes moins familières que ceux des sciences sociales, d'une communication adaptée aux différentes audiences, etc.

4.2. ÉVALUATION SYSTÉMIQUE INCLUSIVE POUR L'ÉGALITÉ DE GENRE, L'ENVIRONNEMENT ET LES VOIX MARGINALISÉES

Les travaux d'élaboration de l'ISE4GEMs (Inclusive Systemic Evaluation fo(u)r Gender equality, Environment and Marginalized voices) ont été initiés par le service d'évaluation indépendant du groupe « Femmes » de l'ONU sur la base du constat suivant : la mise en place d'objectifs interconnectés entre des dimensions environnementale, sociale et économique où le degré d'atteinte des objectifs est dépendant du contexte et échappe en grande partie au contrôle des parties prenantes, nécessite une adaptation du travail évaluatif. L'approche développée s'appuie sur la pensée systémique et la prise en compte de la complexité pour organiser une démarche évaluative participative

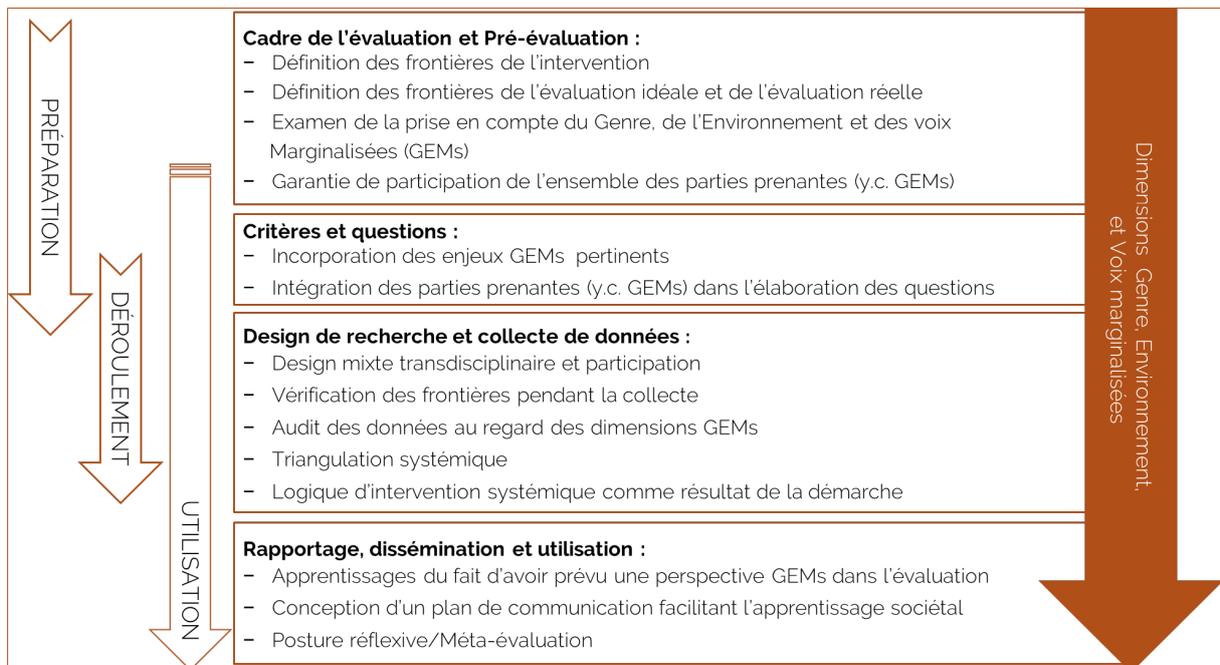
qui met en lumière les liens intersectionnels façonnant la vie des systèmes humain et écologique (Stephens, Lewis et Reddy, 2018). Trois dimensions sont mises en exergue :

- Le genre dans une visée d'égalité de genre ;
- L'environnement afin d'analyser les effets des interventions sur les environnements construits par l'homme (villes, camps de réfugiés, parcs, etc.), sur les systèmes écologiques (forêts, écosystèmes marins, etc.) et les paysages socio-écologiques importants pour le bien-être humain (mines, fermes, champs de pétrole, barrages, etc.) ;
- Les voix marginalisées afin de s'attaquer aux inégalités de pouvoir de certains groupes de personnes en raison de leurs caractéristiques (âge, orientation sexuelle, origine ethnique, religion, handicap, statut, etc.), mais également de certaines voix non humaines telles que la faune, la flore, la culture, etc.

L'examen de chacune de ces dimensions s'accompagne d'une analyse de leurs interconnexions pour sortir des évaluations en silos. L'approche ISE4GEMS a cependant été pensée dans le cadre de l'évaluation de projets de développement et adopte une optique spécifique, celle des évaluations développementales (Patton, cité dans Stephens & al., 2018, p.50). Elle ne s'adapte donc pas directement à l'évaluation d'interventions dans nos pays et/ou à la diversité de la pratique évaluative.

L'approche ISE4GEMS est présentée en quatre phases : (1) Préparation et design (le design devant être compris comme le cadre global de l'évaluation et pas seulement son design de recherche) ; (2) Collecte des données ; (3) Analyse des données et écriture du rapport d'évaluation ; (4) Développement des capacités. Ce découpage, différent de celui que nous avons présenté pour l'approche sensible au genre (cf. Figure 3), ne doit pas cacher la similitude des approches et des éléments qui sont mis en exergue (cf. *infra*). À la figure 6, nous proposons une démarche-type pour une évaluation sensible à la durabilité qui intègre les apports essentiels de l'approche ISE4GEMS tout en pouvant s'appliquer à un large spectre d'évaluations.

Figure 6 : Rendre l'ensemble du processus évaluatif « sensible à la durabilité »



La figure 6 met en exergue la grande similitude entre l'intégration du genre et celle de la durabilité, l'ensemble de la démarche évaluative est impacté. D'une part, la prise en compte de plusieurs dimensions nécessite une vision holistique et une analyse de la complexité de l'intervention et de son évaluation. D'autre part, ces approches s'appuient sur le développement d'une logique d'intervention détaillée dans le cadre d'une démarche inductive et participative. Enfin, une attention particulière est portée à l'utilisation de l'ensemble de la démarche évaluative, son processus et ses livrables, pour soutenir un changement au niveau de l'intervention mais également au niveau sociétal. À la figure 6, certaines étapes ont été regroupées, d'une part, celles de l'établissement du cadre de l'évaluation et de la pré-évaluation et, d'autre part, celles du design de recherche et de la collecte de données. Ces regroupements sont le symptôme de la gêne à présenter de façon linéaire un processus dont le caractère est indiscutablement itératif, *a fortiori* celui d'évaluations participatives adoptant une pensée systémique et complexe.

Dans la suite de la section, nous ne présentons pas l'approche *in extenso* avec tous ses éléments constitutifs (il y aurait trop de redites par rapport aux sections précédentes), mais nous développons les apports essentiels de l'approche ISE4GEMs au niveau de la préparation, du déroulement et de l'utilisation de l'évaluation.

4.2.1 Préparation et définition des frontières de l'intervention et de son évaluation

L'approche s'appuie, à l'image des évaluations sensibles au genre, sur l'analyse systémique de l'intervention, mais elle met également un accent majeur sur l'analyse du système de l'évaluation et sur les relations entre les deux systèmes. La figure 6 regroupe l'établissement du cadre de l'évaluation et de la pré-évaluation montrant ainsi le besoin de connecter les deux systèmes. Même si une demande d'évaluation est adressée par un commanditaire, les frontières du système de l'évaluation ne seront effectivement définies qu'à l'issue d'un processus et celles-ci pourront même être réexaminées en cours d'évaluation, pour faire face à des éléments émergents de la collecte d'informations sur le terrain ou de nouvelles perspectives des parties prenantes.

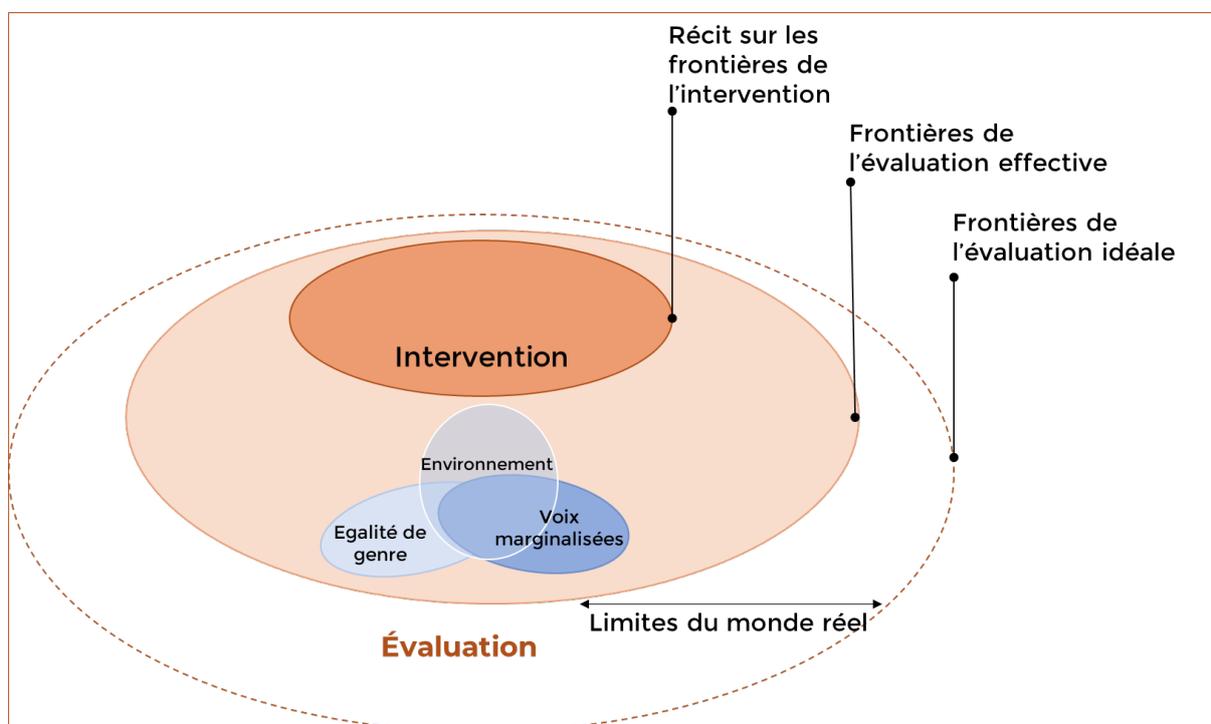
L'établissement des frontières de l'intervention nécessite de considérer les composantes du système (personnes, actions, objectifs, etc.), son contexte (temps, géographie, culture) et les liens avec d'autres systèmes. Cette analyse statique sert d'appui à une compréhension plus dynamique de la manière dont le système s'est créé ou a évolué : Comment cette intervention a-t-elle été élaborée ? Comment a-t-elle été distinguée d'autres systèmes ? Qui fait partie ou pas de ce système ? Qui prend les décisions ? etc. La forme prise par cette analyse est un récit sur ces frontières et sur les changements qui peuvent s'être présentés ou qui ont été mis en œuvre au cours du temps. Durant cette première étape, le questionnement sur les dimensions GEMs se limitera essentiellement à leur inclusion dans l'élaboration de l'intervention, pour quelles raisons, et sur leurs éventuelles connexions avec les autres dimensions. Par exemple, sur la dimension environnementale, les questions suivantes peuvent guider l'analyse :

- L'intervention identifie-t-elle et s'adresse-t-elle à des enjeux environnementaux ? Comment ont-ils été identifiés ? Par qui ? Quelles actions ont été prises ?
- L'intervention identifie-t-elle et s'adresse-t-elle aux relations entre l'environnement et les autres dimensions du genre, de l'économie et des voix marginalisées ? Comment ces relations ont-elles été identifiées ? Par qui ? Quelles actions ont été prises ?

L'analyse s'appuiera dans un premier temps sur les documents stratégiques ou de mise en œuvre de l'intervention, sur un entretien avec le commanditaire de l'évaluation et les acteurs majeurs de la mise en œuvre, sur l'éventuelle logique d'intervention, sur des données de suivi, etc. Cette première étape permet également d'identifier les différentes parties prenantes, leur(s) rôle(s), leurs

interrelations et d'éventuels déséquilibres dans les rapports de force. L'engagement de ces parties prenantes dans une première révision des frontières de l'intervention est un atout (cf. *infra*).

Figure 7 : Les frontières des différents systèmes



Source : Stephens, Lewis et Reddy, 2018 : 61 (notre traduction)

La figure 7 part des frontières de l'intervention et illustre le processus pour aboutir au projet d'évaluation. Comme évoqué, les dimensions du GEMS ne seront pas systématiquement représentées au niveau de l'intervention, un apport de l'approche ISE4GEMS est d'imposer leur prise en considération dans le système de l'évaluation. L'analyse de ce système se déroule en deux étapes : la première vise à établir les frontières de l'évaluation idéale et la seconde, les frontières de l'évaluation effective.

Pour définir l'évaluation idéale, l'intervention va être scrutée en regard des dimensions GEMS. Le questionnement de l'équipe évaluative va porter sur d'éventuels manques dans le récit sur les frontières de l'intervention et sur l'existence de systèmes emboîté(s) ou interconnecté(s) avec celui de l'intervention. Il va également, pour chacune des dimensions GEMS : (1) interroger la manière dont l'évaluation va améliorer la responsabilisation et l'apprentissage quant à cette dimension ; (2) identifier les éléments relatifs à des résultats attendus ou non par rapport à cette dimension ; (3) analyser comment l'évaluation sera utilisée pour améliorer l'intervention et promouvoir la prise en compte de cette dimension. Ce questionnement permet d'identifier les dimensions à intégrer dans l'évaluation. Il est important de noter que toutes les dimensions n'ont pas la même pertinence selon l'intervention, la période d'évaluation et le contexte. Il est donc essentiel d'identifier pour qui l'une ou l'autre des dimensions est primordiale, si l'inclusion d'une dimension dans l'évaluation ne va pas entraver la prise en compte d'une autre en raison des contraintes de ressources, et les implications possibles de ces choix en termes de rapport de forces et d'oppression (Stephens, Lewis et Reddy, 2018, p. 36). La participation des parties prenantes représentant tous les acteurs impliqués et des experts sur les différentes dimensions est à envisager dès cette étape. La faisabilité d'entendre les voix marginalisées humaines et non humaines sera analysée et les adaptations nécessaires étudiées. La définition des frontières de l'évaluation effective demande la prise en considération de la vulnérabilité potentielle de certaines parties prenantes (accès libre au dispositif évaluatif, risque de

blessures liées à la participation, etc.), ainsi que des contraintes du « réel » que ce soit en termes de budget, de temps, d'existence ou d'accès aux données relatives aux dimensions GEMs, etc. Cette double définition met en exergue les limites du travail évaluatif, la distance entre ce qu'il serait pertinent d'évaluer et ce qui l'est. Cette reconnaissance permet de mieux qualifier les résultats, conclusions et recommandations qui sont émis (Stephens, Lewis et Reddy, 2018 : 58).

4.2.2. *Déroulement et flexibilité du design et de la collecte de données*

L'approche ISE4GEMs promeut l'usage de méthodes mixtes et transdisciplinaires qui permet simultanément d'analyser et d'interpréter les interrelations entre les trois dimensions GEMs. Les rapports et dynamiques de pouvoir structurels et interrelationnels sont scrutés. Une place centrale est également occupée par l'engagement participatif pour garantir la pluralité des savoirs, points de vue, perceptions, valeurs et préférences et l'appropriation des problèmes et des solutions. La production de savoir est sociale, partagée et la pratique réflexive est centrale, l'évaluateur bouge entre ses rôles d'expert et de facilitateur.

Le design de recherche et la collecte de données proprement dite sont conçus de façon à pouvoir s'adapter à l'incertitude de la collecte d'informations sur le terrain et aux éléments émergents en cours de processus : la prise en compte d'une nouvelle partie prenante, l'analyse d'un effet inattendu constaté en cours de route, une qualité de données insuffisante sur certaines dimensions GEMs, le besoin d'approfondir les résultats relatifs à un sous-groupe de personnes, etc. Des contrôles réguliers des frontières de l'évaluation idéale et de l'évaluation effective sont effectués. Une attention particulière est apportée à la relation qui a été construite avec les parties prenantes à l'évaluation : le niveau de confiance permet-il l'accès à de nouvelles données ou permet-il à certaines parties prenantes de prendre un rôle plus actif dans l'évaluation ? Par exemple, l'approche ISE4GEMs s'accorde sur l'intérêt d'inclure dans l'équipe d'évaluation un ou une chargé(e) d'évaluation interne à l'organisme de mise en œuvre ou du moins une personne avec une proximité avec l'objet d'évaluation (Stephens, Lewis et Reddy, 2018 : 83) qui puisse amener sa connaissance du contexte et/ou qui puisse acquérir pendant le processus des compétences en évaluation mobilisables ultérieurement (cf. *infra*). L'équipe évaluative révisé également sa perception des rapports de pouvoir entre parties prenantes et veille à les contrer. Des audits des données sont également réalisés pouvant aboutir à une révision du design. Pour permettre ces différentes formes de flexibilité, il est indispensable de prévoir un tampon en termes de ressources, y compris temporelles.

4.2.3. *Déroulement et triangulation systémique*

La phase d'analyse entamée en cours de collecte de données prend un nouveau départ une fois l'ensemble des données rassemblées et codées selon les dimensions et questions évaluatives et selon les dimensions GEMs. L'approche ISE4GEMs préconise d'identifier les forces et les faiblesses au niveau de chacune des dimensions et d'analyser ensuite les interconnexions entre elles. Par exemple, un constat d'accès différencié à des formations professionnelles entre hommes et femmes (dimension évaluative du recours à l'intervention et dimension genre) pourra être connecté à une explication émise par les femmes relative au manque de transports publics rapides (dimension environnement) qui leur permettraient de ne pas s'absenter trop longtemps de chez elles (dimension genre).

Pour assurer la crédibilité et l'appropriation des résultats de l'analyse et des recommandations de l'évaluation, l'approche se fonde sur la triangulation systémique. Celle-ci s'opère en trois étapes : (1) identifier les faits, les éléments de preuve quant aux résultats et aux changements ; (2) permettre aux différentes parties prenantes de donner leur interprétation de ces résultats et de ces changements en fonction de leurs valeurs ; (3) analyser ces résultats et ces changements dans une perspective systémique, c'est-à-dire en prenant en compte leur sensibilité à la définition actuelle des

frontières de l'évaluation. Les résultats seraient-ils les mêmes si les frontières temporelles pour observer les résultats étaient élargies, si les frontières géographiques de l'intervention étaient déplacées dans un autre lieu, si un changement dans un système interconnecté à celui de l'intervention avait lieu, etc. ? Cette distanciation par rapport aux frontières actuelles de l'intervention est cruciale dans l'approche ISE4GEMs, elle permet de fournir des conclusions et recommandations plus nuancées et précises, qui prennent en compte la complexité des processus de changement ainsi qu'une meilleure compréhension des arbitrages à réaliser.

L'intégration d'une perspective de développement durable dans l'évaluation de politiques publiques, tout comme celle du genre, oblige de se distancier du confort des cadres linéaires de « cause à effet » et des pratiques d'évaluation « traditionnelles » centrées sur une logique d'intervention linéaire pour développer des résultats, des conclusions et des recommandations (Bamberger et Segone, 2011). Les évaluations basées sur la théorie peuvent répondre à cet écueil. Une logique d'intervention comme la théorie du changement systémique favorise la mobilisation d'une pluralité de théories et de pratiques réunies au sein d'un même cadre évaluatif, lui-même évolutif (cf. *supra*). Comparativement à une logique d'intervention « classique » (pour peu qu'elle existe), la théorie du changement systémique peut se déconnecter de la façon dont l'intervention sous-revue est censée produire le changement (une vision) pour se focaliser sur le changement social, mais en l'analysant sous différentes perspectives. Elle peut ainsi servir d'outil d'apprentissage et être constamment adaptée et révisée par diverses sources au-delà du travail évaluatif (Stephens, Lewis et Reddy, 2018 : 104).

4.2.4. Utilisation et changement culturel

Le développement des capacités est au cœur de l'approche ISE4GEMs, il en constitue un des buts ultimes. Même si cet objectif prend sans doute plus d'ampleur dans le cadre des politiques de développement, il garde toute sa pertinence dans une région développée comme la Wallonie. C'est « *l'empowerment* » au sens de l'habilitation, « rendre quelqu'un habile à... », plutôt que de l'émancipation (cf. Bélanger, 2011). Ces habiletés consistent à donner plus de visibilité aux inégalités de genre, aux vulnérabilités des groupes marginalisés et aux enjeux environnementaux, de forcer leur prise en compte dans la décision politique, de permettre la participation des trois acteurs - les femmes, les voix marginalisées et l'environnement (en tant qu'acteur) - aux processus de conception, de mise en œuvre et d'évaluation des politiques publiques. On peut généraliser au développement durable la maxime des études de genre, il ne suffit pas de « *just add women and stir* ». En effet, faire du genre ou faire du développement implique nécessairement une sensibilité culturelle et une prise en compte des rapports de force inégaux entre acteurs (y compris l'environnement).

Pour ce faire, l'approche ISE4GEMs a souligné l'importance d'engager les parties prenantes à l'évaluation dans un dialogue à chaque phase de l'évaluation. L'équipe d'évaluation doit en effet veiller à donner une place centrale à la pratique réflexive dans l'ensemble du processus d'évaluation, combinant des rôles d'expert et de facilitateur. L'évaluateur est ainsi responsable de rendre visibles les différentes perceptions de la réalité et les facteurs qui les influencent afin d'examiner de manière critique les conséquences d'adopter une version de la réalité plutôt qu'une autre. Ceci implique d'encourager les parties prenantes à réexaminer de manière critique leurs propres représentations et hypothèses concernant l'intervention publique, les groupes qu'elle vise à atteindre et sur les changements induits par l'intervention. Il s'agit de comprendre collectivement les interactions, dynamiques et modèles qui les sous-tendent. Un processus d'apprentissage mutuel est ainsi amorcé. Avec le temps et la généralisation de cette approche, les différentes parties prenantes acquerront de meilleures compétences pour peser sur la transformation des politiques publiques dans une optique plus durable.

Partant de l'habilitation (ou de l'encapacement) des parties prenantes à l'évaluation, il s'agira de l'étendre à toute la société afin de développer une véritable culture générale de sensibilité aux GEMs. En cela, l'accent dans l'approche ISE4GEMs sur la communication et ses multiples adaptations en fonction des publics à atteindre est primordial pour que toute la société puisse comprendre et porter le changement.

5. Conclusion

L'objet de ce *Working Paper* est de présenter des cadres, approches, outils et méthodes pour intégrer les dimensions du genre et du développement durable dans l'évaluation des politiques publiques. Il vise ainsi à contribuer à l'adaptation de la discipline et de la communauté évaluative aux transformations majeures que traversent notre société et les politiques qui la régissent. Le genre et le développement durable sont deux enjeux incontournables auxquels cette société en mouvance doit faire face. Ils sont transversaux et ne peuvent être cantonnés à certaines phases du cycle de vie des politiques publiques ou à certains domaines politiques. Cette transversalité constitue un réel défi au niveau de leur intégration systématique et généralisée dans la pratique de l'évaluation.

Pour ce qui concerne le genre, sa prise en compte continue à se heurter à un manque de volonté politique. Malgré le fait que le *gender mainstreaming* soit aujourd'hui une obligation légale, les décisions politiques tendent à marginaliser ce sujet en limitant l'intégration d'une approche transversale genre aux seules évaluations d'interventions qui visent spécifiquement à des transformations de genre. La situation est un peu différente pour le développement durable : le cadre international des Objectifs de développement durable dans lequel s'inscrit le Gouvernement wallon constitue une impulsion forte pour pratiquer et généraliser son intégration dans la pratique évaluative.

À côté de la volonté politique, la question du rôle des chargés d'évaluation face à cette intégration de dimensions transversales doit être soulevée. Certains sont peu enclins lors de la définition des frontières de l'évaluation à promouvoir la prise en compte de ces dimensions transversales, sauf si celles-ci sont explicitement incluses dans les objectifs assignés à l'intervention. Par ailleurs, la communauté évaluative manque encore parfois de compétences pour intégrer ces dimensions de manière adéquate. Des solutions se développent pourtant et ce *Working Paper* a tenté de les partager. Les cadres d'analyse de genre tels que les matrices d'analyse de genre et des changements, ainsi que l'approche ISE4GEMs se présentent comme des boîtes à outils concrets et pratiques dépassant largement le niveau assez conceptuel du *gender mainstreaming* ou encore du développement durable. Ils sont un moyen de faciliter la transformation de la pratique des chargés d'évaluation vers une plus grande prise en compte de la complexité des politiques publiques.

Si l'on vise à promouvoir un changement transformateur pérenne de notre société vers davantage d'égalité, d'inclusion et de durabilité, l'adhésion et l'implication de toute la société sont nécessaires. Pour favoriser cette transformation effective du réel, tous les acteurs doivent connaître les informations factuelles, comprendre les mécanismes qui sous-tendent les interventions, percevoir les rapports de pouvoir, etc., autant d'éléments qu'ils puissent mobiliser dans leurs actions quotidiennes. Ce partage de connaissances, cette réflexion critique et ce dialogue sont au centre de l'approche ISE4GEMs (à l'image d'autres approches participatives). C'est ainsi que l'évaluation amorce ou huile les rouages d'un processus d'apprentissage mutuel. Celui-ci devrait s'inscrire dans la durée et s'étendre à toute la société en y développant une véritable culture générale, qui au-delà des enjeux économiques, soit également sensible aux enjeux du genre, de l'inclusion sociale et de l'environnement.

Bibliographie

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 policy-making 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.

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.
- 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.
- 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", *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. Brisolaro, 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.



L'Institut wallon de l'évaluation, de la prospective et de la statistique (IWEPS) est un institut scientifique public d'aide à la prise de décision à destination des pouvoirs publics. Autorité statistique de la Région wallonne, il fait partie, à ce titre, de l'Institut Interfédéral de Statistique (IIS) et de l'Institut des Comptes Nationaux (ICN). Par sa mission scientifique transversale, il met à la disposition des décideurs wallons, des partenaires de la Wallonie et des citoyens, des informations diverses qui vont des indicateurs statistiques aux études en sciences économiques, sociales, politiques et de l'environnement. Par sa mission de conseil stratégique, il participe activement à la promotion et la mise en œuvre d'une culture de l'évaluation et de la prospective en Wallonie.

Plus d'infos : <https://www.iweps.be>



2023