



UNEG
United Nations Evaluation Group

Manuel de l'UNEG pour la conduite de l'évaluation du travail normatif dans le système des Nations Unies

Ce document d'orientation a été préparé par le Groupe de travail de l'UNEG pour l'évaluation du travail normatif, avec le concours de M. Mark Stiles, à qui nous tenons à exprimer notre reconnaissance.

Novembre 2013

Table des matières

Acronymes et abréviations.....	4
Introduction.....	5
Chapitre 1. Déterminer l'évaluabilité du travail normatif	13
Chapitre 2. Préparer l'évaluation	23
Tâche 2.1 : Identifier les parties prenantes et leur rôle dans l'évaluation.....	23
Tâche 2.2 : Déterminer la finalité et les questions clés de l'évaluation.....	25
Tâche 2.3 : Sélectionner les critères autour desquels bâtir l'évaluation	30
Tâche 2.4 : Préparer les indicateurs	33
Tâche 2.5 : Sélectionner la conception et la méthodologie appropriées pour l'évaluation	38
Tâche 2.6 : Sélectionner l'équipe de l'évaluation.....	49
Tâche 2.7 : Préparer le cahier des charges de l'évaluation	50
Chapitre 3. Collecter et analyser les données	54
Tâche 3.1 : Collecter les données	54
Tâche 3.2 : Analyser les données.....	55
Tâche 3.3 : Identifier les enseignements et les pratiques les mieux adaptées.....	58
Chapitre 4. Préparer le rapport d'évaluation.....	62
Tâche 4.1 : Préparer le rapport d'évaluation.....	62
Chapitre 5. Donner suite à l'évaluation	65
Tâche 5.1 : Distribuer le rapport d'évaluation	65
Tâche 5.2 : Préparer une réponse de la part des équipes de direction	66
Tâche 5.3 : Utiliser l'évaluation.....	66
Tâche 5.4 : Évaluer l'évaluation	67
Documents examinés	68

Tableaux

Tableau 1. Exemples de travaux normatifs menés au sein du système des Nations Unies	7
Tableau 2. Étapes de l'appréciation de l'évaluabilité	13
Tableau 3. Défis habituels rencontrés au stade de l'évaluabilité et solutions possibles	15
Tableau 4. Exemple de modèle complété pour catégoriser les parties prenantes	23
Tableau 5. Questions d'évaluation clés par catégorie de travail normatif	28
Tableau 6. Indicateurs pour des activités de plaidoyer de nature normative	34
Tableau 7. Matrice d'évaluation des résultats normatifs liés aux standards en matière de biodiversité.....	37
Tableau 8. Choix des types et conceptions pour les évaluations de travaux normatifs	38
Tableau 9. Méthodes d'évaluation pour les initiatives qui soutiennent l'élaboration d'une norme, d'un standard ou d'une convention international(e)	40
Tableau 10. Méthodes d'évaluation pour les initiatives aidant les gouvernements à intégrer un standard international à la législation, aux politiques générales et aux plans de développement	43
Tableau 11. Méthodes d'évaluation pour les initiatives visant à aider les gouvernements à mettre en œuvre des lois, des politiques et des plans de développement sur la base d'un standard international	44
Tableau 12. Méthodes d'évaluation à utiliser dans des circonstances spéciales	46

Figures

Figure 1. Le cycle de l'évaluation.....	12
Figure 2. Exemple de théorie du changement.....	18
Figure 3. Méthodologies courantes par type d'évaluation.....	45

Encadrés

Encadré 1. Scénario demandant une approche alternative de l'évaluation du travail normatif.	19
Encadré 2. Enseignements des appréciations de l'évaluabilité.....	21
Encadré 3. Contrôle qualité – mener une appréciation de l'évaluabilité.....	21
Encadré 4. Utilisations des évaluations du travail normatif de l'ONU	27
Encadré 5. Critères d'évaluation et questions normatives de l'OCDE/du CAD	31
Encadré 6. Questions pour bâtir l'évaluation d'un programme conçu pour renforcer le soutien institutionnel national apporté aux standards de production plus propre.....	33
Encadré 7. Enseignements tirés de la préparation des évaluations de travaux normatifs	51
Encadré 8. Contrôle qualité – préparer les évaluations de travaux normatifs	52
Encadré 9. Étapes d'une analyse de la contribution	57
Encadré 10. Enseignements en matière de collecte et d'analyse des données.....	60
Encadré 11. Contrôle qualité – collecte et analyse des données.....	60
Encadré 12. Enseignements – Préparer le rapport d'évaluation.....	63
Encadré 13. Contrôle qualité – Préparer le rapport d'évaluation	64
Encadré 14. Enseignements – suivi de l'évaluation.....	67

Acronymes et abréviations

AI	<i>Appreciative Inquiry</i> (Exploration positive)
CEDAW	Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes
PP	Production plus propre
OSC	Organisation de la société civile
FAO	Organisation des Nations Unies pour l'alimentation et l'agriculture
FEM	Fonds pour l'environnement mondial
AIEA	Agence internationale de l'énergie atomique
CRDI	Centre de recherches pour le développement international
OIT	Organisation internationale du travail
OMD	Objectifs du Millénaire pour le développement
CPS	Changement le plus significatif
CAD/OCDE	Comité d'aide au développement/Organisation de la coopération et du développement économiques
CC	Cahier des charges
ONU	Organisation des Nations Unies
PNUD	Programme des Nations Unies pour le développement
UNEG	Groupe des Nations Unies pour l'évaluation
PNUE	Programme des Nations Unies pour l'environnement
UNESCO	Organisation des Nations Unies pour l'éducation, la science et la culture
UNICEF	Fonds des Nations Unies pour l'enfance
ONUDI	Organisation des Nations Unies pour le développement industriel
OMS	Organisation mondiale de la santé

Introduction

Contexte

1. Le Groupe des Nations Unies pour l'évaluation (UNEG – *United Nations Evaluation Group*) est un réseau professionnel qui regroupe les unités chargées de l'évaluation dans le système des Nations Unies et dans les organisations et fonds internationaux qui travaillent en étroite collaboration avec le système des Nations Unies. Il a pour objectif de renforcer l'objectivité, l'efficacité et la visibilité de la fonction d'évaluation et de prôner l'importance de l'évaluation pour l'apprentissage, la prise de décisions et la redevabilité. Dans le cadre de son mandat, l'UNEG a formé un Groupe de travail sur l'évaluation du travail normatif pour répondre à une demande accrue d'évaluations de ce type et au manque de ressources pertinentes.
2. Lors de la réunion générale annuelle de l'UNEG de 2012, le Groupe de travail, après avoir consulté les chefs des équipes d'évaluation des Nations Unies, a défini le travail normatif comme :

Le soutien apporté à l'élaboration de normes et de standards dans les conventions, les déclarations, les cadres de réglementation, les accords, les lignes directrices, les codes de pratiques et autres instruments établissant des standards, aux niveaux mondial, régional et national. Le travail normatif englobe par ailleurs le soutien à la mise en œuvre de ces instruments au niveau des politiques générales, c.-à-d. leur intégration dans la législation, les politiques générales et les plans de développement, et à leur mise en œuvre au niveau des programmes. (UNEG, 2012, p. 5)

3. Il convient de noter que la définition précise trois catégories de travail normatif :
 - a) l'élaboration de normes et de standards ;
 - b) le soutien aux gouvernements et autres entités pour intégrer les normes et les standards à la législation, aux politiques générales et aux plans de développement ; et
 - c) le soutien aux gouvernements et autres entités pour mettre en œuvre la législation, les politiques générales et les plans de développement sur la base de normes, standards et conventions internationaux.
4. Certaines normes, comme les conventions, sont juridiquement contraignantes une fois ratifiées et obligent les États parties à se plier à leurs termes. Par ailleurs, certaines normes comportent des dispositions relatives aux obligations de présentation de rapports par les États parties au sein du système des Nations Unies en ce qui concerne l'application desdites normes¹. Il s'agit, par exemple, de la présentation de rapports sur les cibles

¹ Certains des mécanismes de supervision/de suivi du système des Nations Unies ont été établis il y a longtemps. Par exemple, en ce qui concerne l'OIT, les obligations de présentation de rapports par les États parties ont été établies dans le Traité de Versailles en 1919, et la Commission de l'application des normes de la Conférence internationale du travail et la Commission d'experts pour l'application des conventions ont été fondées dès 1926.

environnementales convenues. D'autres normes peuvent consister en lignes directrices ou instructions générales qui sont applicables sur une base volontaire².

5. Le rôle de leadership de l'ONU dans le travail normatif reste un de ses avantages concurrentiels les plus importants. À travers le consensus, elle établit des normes et des standards universellement applicables pour la paix et la sécurité, le développement économique et socioculturel, les droits de l'homme, l'État de droit, la santé et la durabilité environnementale, entre autres.
6. La vaste majorité des organisations du système des Nations Unies, y compris les fonds, les programmes, les agences spécialisées et les organisations affiliées, entreprennent des travaux normatifs. Pour beaucoup, le travail normatif est au cœur même de leur mandat. Par exemple, les agences spécialisées, comme l'Organisation mondiale de la santé (OMS), l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO), l'Organisation des Nations Unies pour le développement industriel (ONUDI), l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO) et l'Organisation mondiale de la propriété intellectuelle (OMPI), collaborent avec d'autres organisations et États membres de l'ONU pour élaborer des conventions internationales, des standards techniques et des codes de pratiques et pour en soutenir la mise en œuvre. L'Organisation internationale du travail (OIT) travaille avec les États membres pour adopter des normes internationales du travail préparées par les représentants de constituants tripartites (gouvernements, employeurs et travailleurs). D'autres, comme le Fonds des Nations Unies pour l'enfance (UNICEF) et le Programme des Nations Unies pour le développement (PNUD), coopèrent avec les États membres pour parvenir à des standards internationaux en matière de droits de l'enfant, de réduction de la pauvreté, des Objectifs du Millénaire pour le développement (OMD), de la gouvernance démocratique et des interventions en situations de crise³.
7. Le Tableau 1 illustre certaines des formes différentes du travail normatif entrepris par les organisations des Nations Unies. Bien qu'il ne s'agisse pas d'une classification définitive, le tableau donne une idée de la portée et de la variété des travaux normatifs menés au sein du système des Nations Unies.

² Le Manuel de l'UNEG « Intégrer les droits de l'homme et l'égalité des sexes aux évaluations » (2011) propose l'éclaircissement suivant sur l'effet juridiquement contraignant des instruments internationaux relatifs aux droits de l'homme : « *Le droit international des droits de l'homme ne se limite pas aux droits recensés dans les traités. Il englobe également les droits et les libertés qui sont devenus partie intégrante du droit international coutumier, contraignant pour tous les États, y compris ceux qui ne sont pas parties à un traité particulier. Les décisions judiciaires des tribunaux internationaux ou régionaux et des organes internationaux chargés de la surveillance jouent également un rôle considérable dans le droit international des droits de l'homme, car elles apportent des éclaircissements supplémentaires sur la portée des obligations des États et la teneur des droits. Il existe beaucoup d'autres instruments universels et régionaux non contraignants (déclarations, principes, lignes directrices, standards, règles et recommandations) qui ont trait aux droits de l'homme. Ces instruments n'ont aucun effet juridique contraignant, mais ils ont une force morale indéniable et fournissent une ligne de conduite pratique aux États.* »

³ Les organisations internationales et les fonds membres de l'UNEG peuvent aussi jouer un rôle dans le travail normatif, par exemple, dans le cas de l'Organisation internationale de la migration (OIM), pour les questions ayant trait à la migration.

Tableau 1. Exemples de travaux normatifs menés au sein du système des Nations Unies

Types	Description	Exemples
Élaboration de conventions, protocoles et déclarations internationaux	Les conventions constituent le socle des lois internationales, tandis que les protocoles et les déclarations peuvent ne pas être contraignants. Les États membres prennent part à leur élaboration et sont chargés de leur mise en œuvre.	La Convention sur la diversité biologique ; la Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes (CEDAW) ; les conventions de l'Organisation internationale du travail (OIT) concernant les normes du travail et les lois internationales ; la Convention de l'UNESCO concernant la protection du patrimoine mondial culturel et naturel ; la Convention relative aux droits de l'enfant ; la Convention des Nations Unies sur le droit de la mer ; le Règlement sanitaire international et la Convention-cadre pour la lutte anti-tabac (OMS), la Déclaration universelle des droits de l'homme et les Pactes internationaux relatifs aux droits civils et politiques, et aux droits économiques, sociaux et culturels ; et le Protocole de Montréal relatif à des substances qui appauvrissent la couche d'ozone.
Établissement de normes, de standards, de codes de conduite internationaux et de lignes directrices	Nombre d'organisations des Nations Unies prennent part à l'établissement de normes, de standards et de codes de conduite internationaux. Dans certains cas, ils sont adoptés par les États parties sous la forme de législation. Dans d'autres cas, ils peuvent être incorporés dans les politiques générales ou les pratiques de programmes, ou considérés comme des pratiques exemplaires.	CODEX ALIMENTARIUS, le code des standards alimentaires conjoint de l'Organisation mondiale de la santé (OMS) et de l'organisation pour l'alimentation et l'agriculture (FAO) ; le travail de l'Organisation des Nations Unies pour le développement industriel (ONUDI) sur les normes, la métrologie, les essais, l'infrastructure qualité et l'efficacité énergétique ; Le Code de conduite pour une pêche responsable de la FAO ; les standards, réglementations et procédures de l'Organisation de l'aviation civile internationale pour la sûreté, la sécurité, l'efficacité et la régularité de l'aviation ; les standards de sécurité dans le domaine nucléaire de l'Agence internationale de l'énergie atomique (AIEA) ; le Code de conduite de l'ONU concernant l'exploitation et les abus sexuels.
Suivi et présentation de rapports sur la mise en œuvre de conventions, normes et autres obligations	En général, les obligations de présentation de rapports concernent les conventions ratifiées. Cependant, le suivi de l'application des normes par les États parties pourrait aller plus loin.	Le travail inter-agences sur le suivi et la présentation de rapports sur les mécanismes relatifs à la violation des droits de l'enfant dans des situations de conflit armé. Le suivi et la présentation de rapports par le Bureau du Haut-Commissariat aux droits de l'homme (HCDH) sur la mise en œuvre des conventions, des normes et des autres obligations directement sur la base du mandat du Haut-Commissariat. Le travail que mène l'OMS concernant l'éradication et l'élimination de certaines maladies convenues par l'Assemblée mondiale de la santé, et celui effectué par l'Organisation mondiale du commerce en matière de suivi des politiques, réglementations et pratiques commerciales et de renforcement des capacités commerciales ; et, conformément à la Déclaration de l'OIT relative aux

		principes et droits fondamentaux au travail (1998), l'examen des rapports périodiques soumis par les États parties sur les mesures prises pour la mise en œuvre des conventions et recommandations centrales de l'OIT.
Plaidoyer	Les activités de plaidoyer peuvent prendre de nombreuses formes, y compris le suivi, la présentation de rapports et la promotion des normes et des standards directement par le biais du dialogue et de la communication et indirectement en mobilisant les ONG.	Les activités de plaidoyer de l'UNICEF ayant trait à la Convention relative aux droits de l'enfant. Les activités de plaidoyer menées par l'ONU pour soutenir les OMD. Le travail effectué par l'AIEA pour promouvoir l'utilisation sans risque de l'énergie nucléaire. Le travail de l'ONUDI de promotion des industries vertes.
Élaboration et distribution de produits normatifs, y compris produits liés aux connaissances et ressources d'information	Il s'agit, entre autres, de la publication des normes et des standards, de la production et de la distribution d'informations statistiques et de l'entretien des bases de données et, de plus en plus, de la production de ressources basées sur Internet et d'e-apprentissage visant à soutenir la mise en œuvre des normes et standards.	Presque toutes les agences de l'ONU entreprennent des activités de ce type. Par exemple, les rapports de la FAO sur la situation de la pêche, de la foresterie, de l'agriculture, de la sécurité alimentaire dans le monde ; les rapports annuels de l'OIT sur l'application des normes internationales du travail ; l'Indice de développement humain du PNUD ; le Rapport de suivi sur l'Éducation pour tous de l'UNESCO ; l'Indice des droits de l'enfant de l'UNICEF ; les activités de l'ONUDI sur les statistiques industrielles ; le Bilan environnemental annuel du Programme des Nations Unies pour l'environnement (PNUE) ; les bases de données nucléaires de l'AIEA ; la Classification internationale des maladies de l'OMS, sa liste de médicaments essentiels, ses tableaux de croissance infantile, ou son Rapport sur la santé dans le monde ; et les informations statistiques produites et distribuées par certaines des commissions régionales des Nations Unies.
Promotion et renforcement des capacités pour ce qui est des normes et des standards transversaux	Il s'agit entre autres de divers forums, parfois propres à certaines agences et, dans d'autres cas, englobant plusieurs agences, dont l'objectif est de fournir des lieux de rassemblement neutres pour le dialogue sur les aspects techniques et sur les politiques générales.	Promotion et suivi des progrès effectués vers les OMD ; la Déclaration du millénaire sur la protection des groupes vulnérables et les interventions associées ; la promotion du développement durable, y compris la croissance verte ; le travail de l'UNESCO sur la culture et le développement durable ; la fonction de Forum mondial de l'ONUDI ; la promotion et la mise en œuvre par le système des Nations Unies de l'égalité des sexes et de l'intégration du genre dans toutes les activités, sous la conduite d'ONU femmes.
Facilitation du dialogue et de la coordination intergouvernementaux	L'une des plus grandes forces du système des Nations Unies est son pouvoir fédérateur. De nombreuses organisations de l'ONU facilitent le dialogue intergouvernemental sur les	Réunions et négociations intergouvernementaux ; diverses commissions thématiques, comme la Commission de l'ONU sur le développement durable ; la Commission de statistique, la Commission sur la population et le développement, la Commission du développement social ; et une grande partie des activités

	normes et les standards et coordonnent les activités aux niveaux régional et mondial.	des cinq commissions régionales de l'ONU, et les activités de la Commission de la condition de la femme ; et les réunions des groupes d'experts.
--	---	--

8. Non seulement nombre d'organisations des Nations Unies aident les États membres et les institutions partenaires à adopter des normes et des codes de pratique, mais les organisations des Nations Unies sont elles-mêmes tenues d'adhérer à des standards : droits de l'homme, approche basée sur les droits de l'homme en matière de programmation, égalité des sexes et durabilité environnementale, pour n'en nommer que quelques-uns. On peut donc soutenir que l'évaluation du travail normatif est applicable dans l'ensemble du système des Nations Unies parce que presque toutes les organisations des Nations Unies prennent part à des initiatives normatives d'une manière ou d'une autre.

Objectif du manuel

9. L'objectif de ce manuel est d'orienter les organisations des Nations Unies dans le cadre de la conduite d'évaluations de travaux normatifs. Ses objectifs sont les suivants :
- Fournir une approche intégrée de l'évaluation du travail normatif au sein de l'ONU ;
 - Fournir des conseils méthodologiques pratiques, des exemples concrets et des outils pour mener des évaluations de travaux normatifs ; et
 - Mettre en relief les enseignements pertinents et les pratiques les mieux adaptées de l'évaluation de travaux normatifs à l'intérieur et à l'extérieur du système des Nations Unies.
10. Ce manuel est principalement destiné aux professionnels de l'évaluation au sein des Nations Unies qui sont chargés de la gestion et/ou de la conduite d'évaluations de travaux normatifs. Il peut également être utile aux membres du personnel des Nations Unies chargés de la conception, de la gestion et du suivi de programmes et de projets qui intègrent les normes, standards et codes de pratique ; et aux partenaires de l'ONU dans leurs travaux normatifs, en particulier les gouvernements et les organisations de la société civile (OSC). Le manuel peut également être appliqué dans le cadre de l'évaluation des performances internes des organisations des Nations Unies par rapport aux standards relatifs à l'égalité des sexes, aux droits de l'homme, à l'éthique, à la conduite dans le cadre de l'assistance humanitaire et à la durabilité environnementale.
11. Les lecteurs sont susceptibles de demander : « Pourquoi un autre guide de l'évaluation ? En quoi l'évaluation du travail normatif est-elle différente ? »
12. Comme on l'indique ci-dessus, les évaluations du travail normatif constituent une partie considérable des évaluations menées par les organisations membres de l'UNEG, d'après les résultats d'une récente enquête de l'UNEG. Bien que les évaluations autonomes de travaux normatifs soient moins fréquentes, cela pourrait néanmoins changer une fois que les organisations des Nations Unies se mettront à consacrer une plus grande part de leur temps et de leurs ressources au travail en amont à l'avenir. Les agences qui ont efficacement évalué

leur travail normatif au fil des ans perçoivent par ailleurs la nécessité d'un manuel qui saisit et systématise les pratiques efficaces dans l'intérêt de la gestion des connaissances.

13. L'évaluation du travail normatif comporte souvent des défis. Il est souvent difficile d'établir la causalité⁴ dans le travail normatif à mesure que l'on remonte la chaîne logique⁵, et c'est particulièrement difficile au niveau de l'impact⁶. Bien qu'il y ait peut-être un lien direct entre les actions d'une organisation des Nations Unies et l'adhésion par un gouvernement national à une norme particulière, de nombreux facteurs qui échappent au contrôle de l'organisation entrent en jeu au moment de déterminer la mesure dans laquelle le gouvernement traduit la norme convenue en politiques, législation et actions nationales sur le terrain. Les effets directs et l'impact de la plupart des travaux normatifs dépendent d'un environnement habilitant et d'une foule d'actions complémentaires menées par le gouvernement et par les agences des Nations Unies, entre autres. Bien que cela ne requière pas une approche différente de l'évaluation, ni des méthodes différentes, cela demande un important degré d'attention lors de la planification d'une évaluation, et une rigueur renforcée dans l'analyse de la causalité.
14. Une grande partie du travail normatif effectué dans le cadre du système des Nations Unies est difficile à évaluer parce que ses tactiques, produits et effets sont en constante évolution et ne sont pas toujours bien documentés. C'est souvent le cas lors du développement d'un consensus sur un standard mondial et lors du renforcement de la volonté politique pour la mise en œuvre d'une norme internationale. Une grande partie de ce travail n'est pas linéaire et ne cesse d'évoluer. Si l'on peut soutenir que nombre de programmes des Nations Unies sont soumis à des variables contextuelles imprévisibles, une grande partie du travail normatif des Nations Unies rend cette possibilité bien plus concrète.
15. Qui plus est, les travaux normatifs, comme ceux qui portent sur les droits de l'homme ou la durabilité environnementale, tardent à avoir un impact ; il est rare que cela se manifeste dans le calendrier du cycle de programme d'une organisation. Il manque souvent des données concrètes lorsque les organisations des Nations Unies assument un rôle fédérateur ou de coordination, ou bien lorsqu'elles préconisent l'adoption de protocoles, lignes directrices et codes de pratiques internationaux. L'une des raisons de cet état de fait est que les effets de l'intervention⁷ sont souvent clairement stipulés sous forme d'un modèle logique ou d'une théorie du changement⁸.

⁴ La causalité fait référence à la relation entre divers événements, lorsqu'un ensemble d'événements (les effets) sont la conséquence directe d'un autre ensemble d'événements (les causes).

⁵ Une chaîne logique (ou chaîne de résultats ou encore modèle logique) est un modèle visuel qui met en évidence les intrants, les activités, les produits, les résultats et l'impact, et leurs relations causales séquentielles.

⁶ Les impacts sont définis comme les « effets à long terme, positifs et négatifs, primaires et secondaires, induits par une action de développement, directement ou non, intentionnellement ou non » (Comité d'aide au développement de l'Organisation de la coopération et du développement économiques [CAD/OCDE], 2002, p. 24). Les impacts sont généralement mesurés quelque temps après la fin du projet ou du programme.

⁷ Les résultats (ou effets) sont « les effets probables ou atteints à court et moyen termes des produits d'une intervention » (CAD/OCDE, 2002, p. 28). Le terme « résultat » est utilisé de façon interchangeable avec

16. Enfin, souvent il n’y a pas de données précises disponibles pour évaluer le travail normatif, en particulier quand ce dernier n’est pas l’axe principal de l’intervention. Parfois, les membres du personnel de l’ONU eux-mêmes ne sont pas conscients des standards ou ne les ont pas en tête pendant qu’ils effectuent leur travail⁹ ; ils accordent quelquefois une attention insuffisante à l’intégration des standards relatifs à l’égalité des sexes et à la durabilité environnementale dans leur travail ; et il est aussi possible que les parties prenantes comprennent parfois différemment le sens de certaines catégories de travail normatif, comme le « plaidoyer », la « mobilisation sociale » et l’« approche basée sur les droits en matière de programmation » (Stiles, 2002). Par ailleurs, le travail normatif a lieu en grande partie en dehors des projets et des programmes financés grâce à des sources extrabudgétaires, ce qui non seulement contribue à des informations inadéquates, mais peut aussi aboutir à une redevabilité confuse ou trop dispersée (Stiles, 2002). Ces facteurs, entre autres, engendrent des défis méthodologiques profonds pour l’évaluation du travail normatif.
17. Ce manuel s’inspire d’un large éventail de ressources, y compris le manuel *Intégrer les droits de l’homme et l’égalité des sexes aux évaluations* (UNEG, 2011) et les *UNEG Good Practice Guidelines for Follow up to Evaluations* (Lignes directrices relatives aux bonnes pratiques de l’UNEG pour donner suite aux évaluations) (UNEG, 2010a). La terminologie employée dans l’ensemble de ce manuel est conforme aux définitions établies par le Comité d’aide au développement de l’Organisation de la coopération et du développement économiques (CAD/OCDE). Les termes comme « produits » et « impact » sont définis la première fois qu’ils apparaissent.

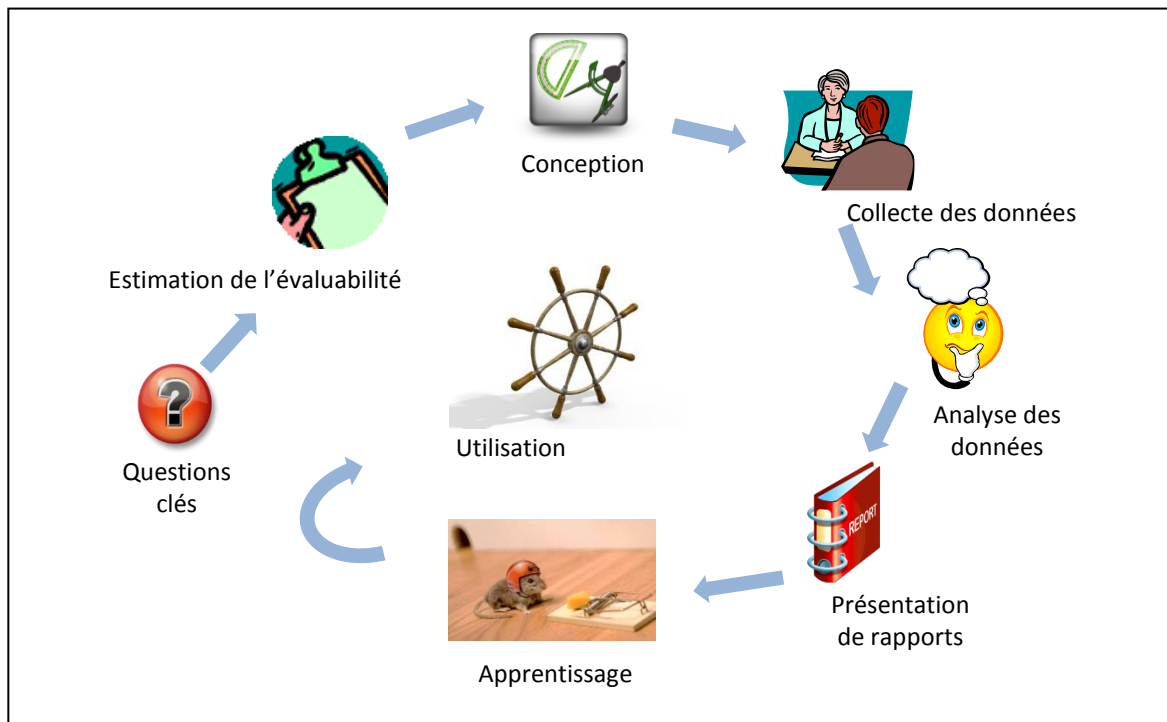
l’expression « effet direct ». Les résultats immédiats ou à court terme devraient être obtenus vers le milieu d’un projet ou d’un programme ; les résultats à court ou moyen terme devraient être obtenus vers la fin.

⁸ La théorie du changement ou théorie du programme est similaire à un modèle logique, mais elle englobe des suppositions clés sous-jacentes aux relations causales et parfois les facteurs principaux (externes et internes à l’intervention) qui sont susceptibles d’influencer les résultats.

⁹ Voir, par exemple, Universal Management Group (2012). *Global Evaluation of the Application of a Human Rights Based Approach to UNICEF Programming*. Résumé d’orientation. New York : UNICEF.

18. Ce manuel est organisé en cinq chapitres, calqués sur le cycle de l'évaluation, illustré par le Figure 1, en commençant par la détermination de l'évaluabilité¹⁰ du travail normatif en question :

Figure 1. Le cycle de l'évaluation



¹⁰ L'évaluabilité est « la mesure selon laquelle une activité ou un programme peut être évalué(e) de manière fiable et crédible » (CAD/OCDE, 2002, p. 21).

Chapitre 1. Déterminer l'évaluabilité du travail normatif

19. Déterminer l'évaluabilité signifie apprécier la situation pour décider si l'évaluation est faisable et financièrement abordable, et si elle présente une valeur suffisante pour être entreprise. Il s'agit entre autres de déterminer si les effets de l'intervention sont correctement définis et vérifiables. Une appréciation de l'évaluabilité est tout particulièrement importante au moment d'envisager l'évaluation du travail normatif à cause des défis méthodologiques en présence, comme on l'évoque dans le chapitre d'introduction.
20. Le Tableau 2 illustre six étapes pratiques par lesquelles il faut passer pour déterminer l'évaluabilité d'un travail normatif.

Tableau 2. Étapes de l'appréciation de l'évaluabilité

Étapes		Questions	Sources des données
1.	Analyser le contexte de l'initiative	Quel est le contexte politique, économique, social et culturel du travail normatif ? Qui sont les parties prenantes et les bénéficiaires clés ?	Documents de référence Rapports de projet/d'avancement Collègues Web Livres Reportages médiatiques
2.	Identifier la nature du travail normatif relatif à l'initiative	Quelle est la nature du travail normatif ? Dans laquelle des trois catégories de travaux normatifs s'inscrit-il, sur la base de la définition de l'UNEG ? Est-il central ou secondaire à l'initiative ?	Documents de référence Rapports de projet/d'avancement Responsable de l'initiative
3.	Déterminer si le travail normatif figure dans la conception de l'initiative, en particulier dans le modèle logique, la théorie du changement, le cadre de développement du programme et/ou le cadre de mesure des performances ¹¹ . S'il n'existe pas de théorie du changement, en développer une et la vérifier	Y a-t-il des effets clairs et réalistes pour le travail normatif aux divers niveaux ? Les suppositions sont-elles clairement formulées dans la logique de programmation ? Y a-t-il des indicateurs valides pour chaque résultat normatif ?	Documents de référence (documents de projet/de programme, y compris modèle logique, etc.)

¹¹ Un cadre de mesure des performances présente, sous forme de matrice, les effets, indicateurs, sources de données, méthodes de collecte des données, fréquence de la collecte et personnes ou organisations chargées de la collecte des données prévus pour l'intervention. La matrice sert de base pour un plan de suivi détaillé.

	avec les responsables de l'intervention ¹² .		
4.	Déterminer s'il y a des données robustes et facilement disponibles ayant trait au travail normatif	<p>Quelles sont les données de suivi qui existent ? Quelles en sont la qualité et la couverture ?</p> <p>Quelles sont les informations de référence disponibles ?</p> <p>Où peut-on se procurer des informations narratives et anecdotiques ?</p>	<p>Documents de référence</p> <p>Rapports de projet/d'avancement</p> <p>Collègues</p> <p>Responsable/spécialistes</p>
5.	Déterminer l'utilisation de l'évaluation	<p>Quelle sera la finalité de l'évaluation ?</p> <p>Comment sera-t-elle utilisée ?</p> <p>Quels intérêts l'évaluation revêt-elle pour les parties prenantes ? Quelles sont leurs attentes et leurs besoins en matière d'information ?</p> <p>Qui bénéficiera de l'évaluation ?</p> <p>Dans quelle mesure les utilisateurs primaires de l'évaluation sont-ils réceptifs à l'apprentissage et au changement ?</p>	<p>Équipe de direction supérieure</p> <p>Spécialistes</p> <p>Partenaires et parties prenantes clés</p>
6.	Estimer en gros le temps et les ressources requis pour l'évaluation et déterminer s'ils sont disponibles	<p>Pour quand l'évaluation doit-elle être effectuée, p. ex. pour la prise de décisions ou l'établissement des budgets ?</p> <p>Quelles sont les connaissances techniques spécialisées requises pour l'évaluation ?</p> <p>Quel sera le coût de l'évaluation ?</p> <p>Les ressources requises – argent et connaissances techniques spécialisées – sont-elles disponibles et abordables ?</p>	<p>Collègues chargés de l'évaluation</p> <p>Documents sur des évaluations similaires</p>

21. La quatrième étape – déterminer s'il y a des données suffisantes pour pouvoir continuer – est pertinente pour les évaluations de toutes les formes de travaux normatifs et tout particulièrement importante pour évaluer le soutien apporté par l'ONU à l'élaboration de normes et de standards. Les interlocuteurs (sources d'information) clés – personnes dotées d'informations historiques riches sur le rôle de l'ONU – ont peut-être changé au fil du temps et il est possible qu'il n'y ait guère de données factuelles documentées ayant trait à la nature et à la qualité du travail effectué par l'ONU. Dans ces cas, le responsable de l'évaluation doit

¹² Ceci est un guide complet récent pour la préparation des théories du changement : Funnell, S. et Rogers, P. (2011). *Purposeful Program Theory*. San Francisco : Jossey-Bass.

soit concentrer l'évaluation sur des domaines où il y a suffisamment de données, soit recommander de ne pas mener à bien l'évaluation.

22. La cinquième étape – déterminer l'utilisation des informations de l'évaluation – est tout particulièrement importante. Si l'évaluation du travail normatif a peu de finalité pratique, si les parties prenantes clés ne se montrent pas intéressées par la perspective d'en tirer les enseignements, et s'il y a des occasions limitées d'entraîner des changements positifs, alors il n'est pas sûr que l'évaluation doive avoir lieu. Cette étape se fonde avec les tâches décrites dans le chapitre suivant, comme se procurer les points de vue des parties prenantes sur les questions les plus importantes pour l'évaluation, déterminer comment l'évaluation devrait être conçue, quand elle devrait avoir lieu et comment les informations générées peuvent être utilisées pour éclairer les décisions et améliorer les performances. Les informations de ce type nous aideront à axer l'évaluation sur les questions les plus pertinentes pour ses utilisateurs.
23. Il peut aussi se révéler utile de prévoir une phase de lancement dans le cadre de l'appréciation de l'évaluabilité, en particulier pour les évaluations complexes d'envergure, comme les évaluations mondiales, régionales, les évaluations de pays et les évaluations thématiques de travaux normatifs. Une mission de lancement peut parfois contribuer à renforcer l'appropriation de l'évaluation par ses utilisateurs prévus, et elle peut résoudre certains des défis habituels auxquels se heurtent les évaluateurs à ce stade, comme on le décrit dans le Tableau 3.

Tableau 3. Défis habituels rencontrés au stade de l'évaluabilité et solutions possibles

	Défis	Solutions possibles
1.	Le contexte de l'intervention est complexe.	<ul style="list-style-type: none"> • Prévoir assez de temps et de ressources pour comprendre le contexte. • Entreprendre une recherche sur Internet et/ou une recherche documentaire. • Entreprendre une mission de lancement. • Prévoir de recruter des experts locaux capables de combler les lacunes contextuelles.
2.	Le travail normatif semble être un élément secondaire de l'intervention.	<ul style="list-style-type: none"> • Tenter de comprendre les raisons pour lesquelles le travail normatif pourrait être secondaire en contactant les personnes chargées de la conception et de la mise en œuvre de l'intervention. • Déterminer si l'évaluation du travail normatif revêt une valeur, p. ex. pour éclairer la programmation future, dans le cadre de laquelle le travail normatif aura un rôle plus central.
3.	Le travail normatif ne se reflète pas dans un modèle logique, une théorie du changement ou un cadre de programme de pays.	<ul style="list-style-type: none"> • Tenter de déterminer pourquoi le travail normatif ne s'y reflète pas en contactant les personnes chargées de la conception et de la mise en œuvre de l'intervention. • Construire une théorie du changement rétrospective qui englobe le travail normatif et la valider avec les personnes chargées de l'intervention.
4.	Le travail normatif se reflète	<ul style="list-style-type: none"> • Si les défauts sont considérables, comme le fait que nombre des

	dans le modèle logique, la théorie du changement ou le cadre de programme de pays, mais ces derniers présentent de sérieux défauts.	effets, des indicateurs et des suppositions sont peu réalistes, mener, si possible, une négociation avec les parties prenantes du programme et les revoir.
5.	Les documents ne fournissent guère d'informations sur le travail normatif et ses effets.	<ul style="list-style-type: none"> • Contacter les personnes chargées de la mise en œuvre pour comprendre la nature du travail normatif. Se procurer des informations auprès d'autres sources, comme les gouvernements nationaux, les rapports d'autres organes onusiens, des organisations de développement et des ONG locales. • Se concentrer sur les effets prévus sur lesquels l'organisation chargée de la mise en œuvre¹³ exerçait un contrôle (au lieu des impacts de haut niveau).
6.	Il n'y pas assez de temps ou de ressources pour évaluer complètement le travail normatif.	<ul style="list-style-type: none"> • Réduire l'échelle de l'évaluation et l'axer sur les résultats normatifs à court et moyen terme, plutôt que sur les impacts normatifs. • Chercher des manières de gagner du temps et d'économiser des ressources, comme une étude documentaire, l'auto-évaluation ou une ou deux études de cas riches en informations.

« Le cadre de référence du Quatrième bilan était extrêmement ambitieux, au point que la mission d'évaluation n'a pas pu exploiter l'ensemble des données disponibles et des résultats des évaluations dans les délais impartis et dans les limites des budgets alloués aux études annexes. En conséquence, le présent rapport ne peut répondre pleinement à certaines des questions que soulèvent les points importants évoqués plus loin. »

- Fonds pour l'environnement mondial, *Progrès vers la réalisation d'un impact. Quatrième bilan global du FEM*, 2010, p. 4

24. La rectification du troisième problème – l'absence de travail normatif dans le modèle logique, la théorie du changement ou le cadre du programme – doit avoir lieu en consultation avec le responsable du programme et les parties prenantes clés du programme. Dans l'idéal, un tel modèle ou cadre devrait être produit lors d'un atelier bien facilité, avec la pleine participation des parties prenantes clés. Bien qu'un atelier demande du temps et des ressources, il pourrait apporter une plus-value en renforçant les capacités des parties prenantes et en favorisant l'appropriation de l'évaluation par ses utilisateurs prévus.

25. Au moment d'élaborer une théorie du changement, il est important d'identifier le degré de contrôle que l'organisation des Nations Unies dont il s'agit exerce sur les effets à court et moyen terme de l'intervention. Cela permettra d'identifier ultérieurement la portée et l'axe de l'évaluation. Si l'intervention s'inscrit dans la première catégorie de travail normatif – la création de normes, de standards et d'instruments internationaux – alors l'évaluation devrait se concentrer sur les questions de processus et de gouvernance et, éventuellement, sur la

¹³ L'expression « organisation chargée de la mise en œuvre » fait généralement référence à l'organisation des Nations Unies qui prend part à des activités normatives, soit dans le cadre d'un projet ou d'un programme, soit dans le cadre d'initiatives formelles de plaidoyer ou de facilitation, comme l'organisation de réunions et l'élaboration de standards avec des groupes d'experts techniques. Elle peut également faire référence aux travaux normatifs entrepris par une organisation des Nations Unies par l'intermédiaire d'ONG partenaires. Il est entendu qu'il incombe en dernière instance aux gouvernements nationaux, en tant que garants d'obligations, de mettre en œuvre les normes, standards et conventions internationaux.

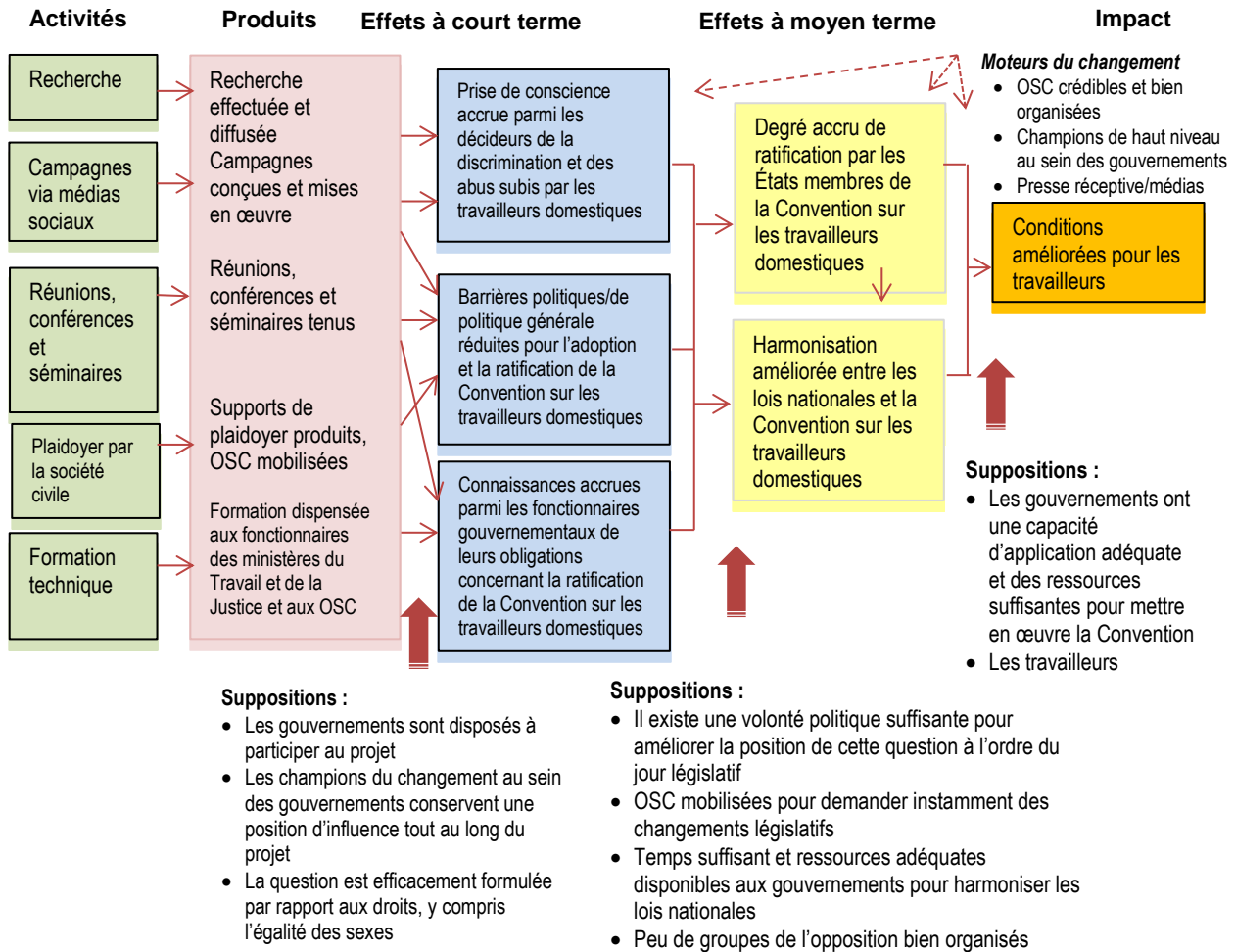
pertinence de la norme, du standard ou de l'instrument international. Si l'intervention a pour objectif d'aider un gouvernement à adopter et à mettre en œuvre une norme particulière en tant que garant d'obligations, alors l'évaluation devrait porter sur les effets liés à ce soutien.

26. Les effets immédiats ou à court terme indiquent généralement des changements au niveau des connaissances, de la prise de conscience et de l'accès aux ressources parmi les bénéficiaires de l'intervention. Les effets intermédiaires ou à moyen terme indiquent en général des changements au niveau du comportement ou des pratiques de particuliers et de groupes et – pour les programmes de longue durée – au niveau des capacités des organisations et des institutions. L'organisation chargée de la mise en œuvre exerce normalement un contrôle et une influence considérables sur ces effets-là et sur ceux à court terme. C'est au niveau des impacts ou des effets à long terme que se présentent les problèmes d'attribution les plus épineux¹⁴. Par exemple, bien qu'il soit réaliste de penser que le travail normatif effectué par une organisation des Nations Unies puisse aboutir à des politiques, des réglementations et des lois renforcées alignées sur des standards internationaux, il serait excessif de penser que cette organisation de l'ONU pourrait contribuer directement à améliorer la vie des populations ciblées, tout au moins dans le cadre des calendriers et budgets habituels des projets ou programmes. De même il serait irréaliste d'attendre d'une organisation des Nations Unies qu'elle obtienne à elle seule des impacts macroéconomiques, comme une hausse des salaires, la réduction de la vulnérabilité des pauvres ou des salaires égaux pour les femmes.
27. Un problème fréquent associé est celui de modèles logiques et de théories de programme qui cherchent à saisir tous les résultats normatifs possibles plutôt que ceux qui sont au cœur même du travail normatif. Le modèle logique ou la théorie du changement devraient se concentrer sur les effets les plus importants à obtenir étant donné les ressources et la portée de l'intervention. Dans l'idéal, la logique de programme d'un projet typique dans le système de l'ONU devrait avoir un énoncé d'impact, deux ou trois effets à moyen terme et entre trois et cinq effets à court terme. Ce qui est important, c'est de bien formuler les principaux résultats ; tout résultat supplémentaire identifié à l'aide du suivi et de l'évaluation est un avantage¹⁵.
28. Un certain nombre de guides – dont certains figurent à la fin de ce chapitre – peuvent aider les praticiens de l'évaluation à préparer des modèles logiques et des théories du changement. La Figure 2 illustre une théorie du changement pour un projet à long terme fictif dont l'objectif est de faire en sorte qu'un nombre accru de pays ratifient la Convention de l'OIT sur les travailleurs domestiques.

¹⁴ L'attribution est la « confirmation d'une relation causale entre les changements observés (ou que l'on s'attend à observer) et une action spécifique ». (CAD/OCDE, 2002, p. 17).

¹⁵ L'évaluation du travail normatif de l'UNESCO dans le domaine de la culture (première partie) – Convention de 2003 pour la sauvegarde du patrimoine culturel immatériel (2013) constitue un bon exemple d'évaluation pour laquelle une théorie du changement rétrospective a été préparée et utilisée tout au long du processus d'évaluation : unesdoc.unesco.org/images/0022/002230/223095E.pdf

Figure 2. Exemple de théorie du changement



29. L'absence de données cruciales – le cinquième problème – pourrait justifier l'ajournement de l'évaluation. Il peut manquer des données parce que l'intervention n'a pas fait l'objet d'un suivi systématique et qu'aucune donnée de suivi sur les progrès effectués et les effets obtenus n'a donc été recueillie. Elle peut également être due au fait que les effets à moyen terme et l'impact, comme par exemple les avantages réels pour les citoyens et les économies nationales, ne se manifesteront probablement pas pendant de nombreuses années, comme l'a conclu une évaluation menée en 2009 par la FAO :

La création d'un instrument international est, au mieux, un résultat provisoire. [T]oute détermination de l'utilité réelle des instruments internationaux doit dépendre des effets et des impacts qui surviennent après la création de ces instruments : entrée en vigueur, opérations internationales, mise en œuvre nationale et application sur le terrain. Ces effets n'apparaîtront en général que progressivement, sur dix ou vingt ans après l'adoption d'un instrument international. (FAO, 2009, pp. 2-3)

30. Autre exemple : une intervention qui englobait une campagne dans les médias de masse n'a pas encore engendré suffisamment de données sur l'utilisation médiatique. Dans ce cas, il peut être judicieux d'ajourner l'évaluation jusqu'à ce que ces données aient été recueillies.
31. Une partie considérable des travaux normatifs de l'ONU a lieu dans des environnements instables où il est difficile d'établir les effets et où la causalité est problématique. Ces conditions peuvent demander des approches non traditionnelles de l'évaluation. Considérons le scénario dans l'Encadré 1.

Encadré 1. Scénario demandant une approche alternative de l'évaluation du travail normatif.

Le pays X a signé mais pas ratifié le Pacte international des Nations Unies relatif aux droits civils et politiques (PIDCP). Cela fait plusieurs années que des organisations de l'ONU et d'autres exercent une pression discrète sur le gouvernement national pour qu'il respecte l'article 27 du PIDCP, qui traite des droits des groupes minoritaires à la liberté culturelle, religieuse et linguistique. Rares sont les activités de ce plaidoyer qui ont été documentées, et la collaboration entre les Nations Unies et les agences donatrices a été sporadique. Les ONG locales et internationales ont aussi mené un plaidoyer en faveur du changement, de façon plus appuyée que l'ONU et les bailleurs de fonds bilatéraux. L'organisation onusienne chef de file impliquée dans ce travail a reçu un financement extrabudgétaire considérable pour ce travail. Elle est dotée d'un modèle logique, mais il est dépassé, et l'organisation n'a pas systématiquement recueilli des données par rapport aux indicateurs. La capacité de gestion axée sur les résultats est faible et la présentation de rapports a largement tourné autour des activités. Le gouvernement national a établi des programmes de développement économique pour les minorités. Il a promis une nouvelle législation, mais est revenu sur cette promesse chaque fois qu'il y a eu des protestations, dont certaines violentes. L'organisation onusienne chef de file négocie actuellement une subvention considérable de la part d'un bailleur de fonds pour une nouvelle initiative dont l'objectif est d'aider le pays à adhérer à l'article 27 et à aligner ses politiques et lois intérieures en conséquence. L'organisation de l'ONU souhaite évaluer sa programmation afin d'en tirer les enseignements et de les appliquer à sa nouvelle initiative.

32. En supposant qu'une appréciation de l'évaluabilité a conclu qu'une évaluation conventionnelle ne revêtirait guère de valeur, quelles sont les autres options possibles ? Les conditions décrites dans le scénario correspondent bien aux critères de l'évaluation développementale. L'évaluation développementale est un type d'évaluation qui permet de *développer* des programmes et des projets, et qui est tout particulièrement adaptée à l'innovation dans les programmes menés dans des environnements complexes (d'où le sous-titre « *Applying Complexity Concepts to Enhance Innovation and Use* » (Appliquer des concepts de complexité pour améliorer l'innovation et l'utilisation) du livre de Michael Patton publié en 2011, *Developmental Evaluation*). Dans l'évaluation développementale, l'évaluateur devient généralement un membre de l'équipe chargée de la mise en œuvre et participe pleinement aux décisions et facilite les discussions sur ce qu'il faut évaluer et comment. L'évaluateur aide les membres de l'équipe à conceptualiser, concevoir et mettre à l'épreuve de nouvelles approches dans le cadre d'un processus continu d'expérimentation et d'adaptation, se montrant toujours très sensible aux résultats inattendus et aux changements survenus dans l'environnement. La fonction primaire de l'évaluateur est d'introduire dans les discussions en équipe une réflexion évaluative, d'apporter des données robustes aux délibérations stratégiques et de faciliter une réflexion et une prise de décisions axées sur des données factuelles et systématiques. En bref, l'évaluation développementale a pour objectif d'aider les programmes menés dans des environnements complexes à bien démarrer et à obtenir de bons résultats jusqu'à leur conclusion à travers un processus d'apprentissage, d'adaptation et d'innovation continus.
33. L'évaluation développementale peut se révéler tout particulièrement applicable à la première catégorie de travail normatif, l'élaboration de normes, standards et conventions internationaux, où elle pourrait fournir des données riches permettant d'éclairer les décisions relatives au processus, à la participation et à la communication. Il est bien évident que, du fait que dans l'évaluation développementale l'évaluateur tend à devenir un membre de l'équipe chargée de la mise en œuvre, ce type d'évaluation n'est pas applicable lorsque l'évaluation est censée être strictement indépendante, comme c'est généralement le cas dans le système des Nations Unies.
34. En définitive, la décision de mener ou non une évaluation – conventionnelle ou non conventionnelle – est affaire de jugement. Il n'y a pas de formule ni d'outils universellement applicables pour déterminer ce qu'il convient de faire.

Encadré 2. Enseignements des appréciations de l'évaluabilité

Une approche participative de l'élaboration d'une théorie du changement peut contribuer à révéler des suppositions et convictions implicites qui, souvent, ne sont pas communiquées aux particuliers et aux groupes prenant part à une initiative.

Préparez-vous à suivre et à évaluer le travail normatif lors de la conception d'un programme en incorporant le travail normatif dans la théorie du changement et le cadre de mesure des performances. Cela vous aidera non seulement à assurer la collecte de données de référence et un flux régulier d'informations sur les performances durant tout le suivi, mais cela vous donnera aussi plus de chances de disposer de bonnes données à temps pour une évaluation.

Une appréciation de l'évaluabilité peut se révéler utile que l'évaluation ait lieu ou non, parce qu'elle précipite souvent des améliorations considérables des conceptions des programmes en mettant en évidence des buts confus, une logique imparfaite, des suppositions peu raisonnables et des attentes contradictoires ou peu réalistes sur ce que le programme a entrepris de faire (Rog, 1985).

C'est lorsqu'elles sont menées en équipe que les appréciations de l'évaluabilité donnent souvent les meilleurs résultats. L'équipe devrait englober l'évaluateur principal et le responsable de l'évaluation, les personnes chargées de la mise en œuvre du programme, les administrateurs et d'autres groupes de parties prenantes (Morra-Imas et, 2009).

Les théories du changement efficaces sont souvent multidimensionnelles. Par exemple, une théorie du changement liée à l'autonomisation des femmes qui ne se concentre que sur une dimension, comme le développement économique et la formation en compétences, tout en ignorant des facteurs comme l'accès aux marchés et les entraves socioculturelles, est insuffisante (UNEG, 2011).

Encadré 3. Contrôle qualité – mener une appréciation de l'évaluabilité

- ✓ L'appréciation de l'évaluabilité a-t-elle été effectuée conformément au processus en six étapes décrit dans le présent manuel ?
- ✓ Les questions ont-elles été posées et y a-t-on répondu à chaque étape ?
- ✓ L'appréciation a-t-elle été effectuée avec la participation de l'évaluateur principal et des parties prenantes clés ?
- ✓ Une théorie du changement a-t-elle été élaborée qui saisit clairement la logique du programme et ses suppositions sous-jacentes ?

Ressources utiles

Dozois, E., Langois, M. et Blanchet-Cohen, N. (2010). *DE.201. A Practitioner's Guide to Developmental Evaluation*. Montréal : The J. W. McConnell Family Foundation. Disponible sur : [<www.mcconnellfoundation.ca/>](http://www.mcconnellfoundation.ca/)

Funnell, S. et Rogers, P. (2011). *Purposeful Program Theory*. San Francisco : Jossey-Bass.

- Gamble, J. (2008). *A Developmental Evaluation Primer*. Montréal : The J. W. McConnell Family Foundation. Disponible sur : <www.mcconnellfoundation.ca/>
- GrantCraft (2006). *Mapping Change: Using Theory of Change to Guide Planning and Evaluation*. Disponible sur : <www.grantcraft.org>
- Patton, M. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York : Guilford Press.
- Reisman, J., Gienapp, A., et Stachowiak, S. (2007). *A Guide to Measuring Advocacy and Policy*. Baltimore : Annie E. Casey Foundation.
- Fonds des Nations Unies pour l'enfance (2010). *Advocacy Toolkit: A guide to influencing decisions that improve children's lives*. New York : UNICEF. Disponible sur : <www.unicef.org/evaluation/files/Advocacy_Toolkit.pdf>
- W. K. Kellogg Foundation (2004). *Using Logic Models to Bring Together Planning, Evaluation, and Action. Logic Model Development Guide*. Battle Creek, Michigan : W.K. Kellogg Foundation. Disponible sur : <www.wkkf.org/knowledge-center/resources/2006/02/wk-kellogg-foundation-logic-model-development-guide.aspx>

Chapitre 2. Préparer l'évaluation

35. Une fois que l'appréciation de l'évaluabilité a été menée à bien et que le responsable de l'évaluation a déterminé que l'évaluation aura lieu, l'étape suivante consiste à préparer le cahier des charges. Ce chapitre décrit les tâches requises, sur la base des enseignements tirés d'évaluations antérieures de travaux normatifs. Les tâches ne sont pas forcément séquentielles : tout dépend des circonstances, des politiques de l'unité chargée de l'évaluation et de l'avis personnel du responsable de l'évaluation.

Tâche 2.1 : Identifier les parties prenantes et leur rôle dans l'évaluation

36. Il est important, à ce stade, de réfléchir à qui devrait prendre part à l'évaluation et à quel moment, et à la manière dont ces personnes devraient participer afin de maximiser l'utilité de l'évaluation. Il est important d'identifier les détenteurs de droits et les garants d'obligations ainsi que leurs rôles respectifs dans l'évaluation, en particulier pour le travail normatif qui soutient l'adoption ou la mise en œuvre de normes et conventions internationales. Il faut commencer par une analyse des parties prenantes, illustrée dans le Tableau 4 ci-dessous, qui se base en gros sur le scénario présenté dans l'Encadré 1 au chapitre précédent. Un aspect important est que l'analyse des parties prenantes identifiera les utilisateurs de l'évaluation comme étant distincts des bénéficiaires de l'intervention faisant l'objet de l'évaluation.

Tableau 4. Exemple de modèle complété pour catégoriser les parties prenantes

Bénéficiaires de l'intervention	Utilisateurs (clients) de l'évaluation
<p><i>Primaires (directs/indirects)</i></p> <p>Homologues gouvernementaux (directs) Groupes représentant les minorités (directs)</p> <p>Minorités (indirects)</p>	<p><i>Primaires</i></p> <p>Responsables locaux de l'organisation des Nations Unies chargée de la mise en œuvre Responsables supérieurs de l'organisation des Nations Unies chargée de la mise en œuvre (siège et région) Homologues gouvernementaux</p>
<p><i>Secondaires</i></p> <p>Fonctionnaires haut-placés (directs) Spécialistes culturels des minorités, experts en linguistique et chefs religieux (directs et indirects)</p>	<p><i>Secondaires</i></p> <p>Bailleurs de fonds collaborateurs ONG collaboratrices Autres organisations des Nations Unies Groupes représentant des minorités Universitaires et professionnels de l'évaluation</p>
<p><i>Autres</i></p> <p>Autres groupes minoritaires ou fonctionnaires gouvernementaux non touchés par l'intervention</p>	

37. Les différentes catégories de travaux normatifs auront vraisemblablement des parties prenantes différentes. Par exemple, l'élaboration de normes et standards fait souvent intervenir des groupes de travail techniques issus des pays participants, avec le soutien administratif de l'ONU. Les travaux normatifs qui englobent un soutien aux

gouvernements nationaux en tant que garants d'obligations pour les aider à ratifier, intégrer et mettre en œuvre les normes font souvent intervenir un éventail beaucoup plus large de parties prenantes, comme des départements techniques gouvernementaux, des OSC, des instituts de recherche et des organismes techniques. Par exemple, l'évaluation de 2010 par l'UNESCO de son Objectif stratégique de programme 11, « Protéger et promouvoir de façon durable le patrimoine culturel », a fait intervenir plus de 100 membres du personnel et experts externes et presque 200 bénéficiaires et partenaires chargés de la mise en œuvre. L'évaluation de 2009 menée par l'UNICEF de DevInfo, un système de base de données conçu pour suivre le développement humain, a fait intervenir des experts internes et externes, ainsi qu'une ample gamme d'utilisateurs du système dans six pays.

38. Au moment d'évaluer des travaux normatifs auxquels participent de nombreuses parties prenantes, un exercice de cartographie constitue un outil utile, car il produit une représentation visuelle des parties prenantes et des relations entre elles. Une carte des parties prenantes peut révéler des informations contextuelles importantes, ainsi que des questions qu'il vaut la peine d'examiner. Dans une récente évaluation menée par la FAO, l'exercice de cartographie a aussi été utilisé pour recueillir les perceptions d'une large gamme de parties prenantes sur la pertinence et l'efficacité des activités de la FAO portant sur le régime foncier, les droits et les questions d'accès¹⁶. Une théorie du changement peut se révéler utile pour identifier les parties prenantes par rapport aux processus et aux résultats à différents niveaux.
39. Les responsables des évaluations devraient inclure dans l'analyse des parties prenantes les personnes qui ont peut-être été exclues de la participation au travail normatif, ou qui pourraient avoir subi les effets négatifs de ce travail. Ils doivent aussi veiller à ce que les points de vue des femmes ainsi que des hommes soient bien représentés au moment de décider qui devrait participer à une évaluation du travail normatif. Cela ne signifie pas forcément qu'il doit y avoir un nombre égal de femmes et d'hommes ; en revanche, cela signifie que la sélection devrait permettre aux voix et aux points de vue des femmes ainsi que des hommes de se faire entendre.
40. La question de savoir s'il convient de sélectionner dans la colonne de gauche, celle de droite ou les deux dans le Tableau 4 dépend de la finalité de l'évaluation et du degré de participation requis. À une extrémité du continuum participatif se trouve l'évaluation de l'autonomisation, dans le cadre de laquelle l'évaluateur transmet le contrôle de la prise de décisions, la collecte de données et l'analyse des données aux bénéficiaires primaires du programme (Fetterman, 2001). Bien qu'il soit difficile de bien l'effectuer, ce modèle d'évaluation est applicable dans des situations où le renforcement de la capacité d'évaluation des parties prenantes, comme par exemple les détenteurs de droits, est un objectif premier. À l'autre extrémité du continuum, on trouve l'évaluation indépendante

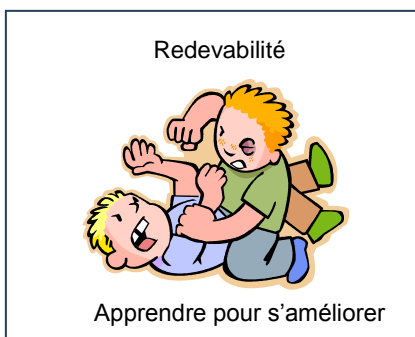
¹⁶ Évaluation stratégique des activités de la FAO en matière de régimes fonciers et de droits applicables à la terre et aux autres ressources naturelles et sur la question de l'accès (janvier 2012). Annexe 3 : Étude des perceptions des parties prenantes. Disponible sur : www.fao.org/docrep/meeting/025/mC957E.pdf

par un tiers. Dans ce cas, l'évaluateur externe exerce un contrôle complet, sans guère de participation de la part des personnes intervenant dans l'initiative normative. Cette approche – commune à nombre d'unités d'évaluation des Nations Unies – est tout particulièrement adaptée dans les situations où il peut y avoir une controverse ou des enjeux importants qui nécessitent l'indépendance des évaluateurs et où la redevabilité est le but clé.

41. Pour de nombreuses évaluations entreprises dans le système des Nations Unies, le modèle de prédilection se situe quelque part entre ces deux extrêmes et la participation des utilisateurs prévus (colonne de droite du Tableau 4) est essentielle pour maximiser l'utilisation de l'évaluation. Les principes de la Déclaration de Paris relatifs à l'« appropriation » et à la « responsabilité mutuelle » renforcent ce choix, et indiquent que les évaluations devraient être menées conjointement avec les gouvernements ou, tout au moins, avec la participation active des gouvernements.
42. Un mécanisme permettant d'améliorer la participation des utilisateurs, qui est pertinent pour chacune des trois catégories du travail normatif, est le groupe de référence de l'évaluation. Ce manuel ne traite pas de la manière de mettre en place ou du fonctionnement d'un groupe de référence de l'évaluation, car il existe d'autres manuels plus généraux sur l'évaluation qui donnent de nombreux conseils sur cette question.
43. La mesure dans laquelle le responsable de l'évaluation fait intervenir les parties prenantes dans la collecte et l'interprétation des données dépend du caractère sensible des questions couvertes par l'évaluation, de la façon dont l'évaluation sera utilisée et du temps et des ressources disponibles. Il n'y a pas de règles fixées. Il est parfois nécessaire de renoncer à un peu de rigueur méthodologique pour maximiser la participation.

Tâche 2.2 : Déterminer la finalité et les questions clés de l'évaluation

44. Déterminer la finalité de l'évaluation est peut-être la tâche la plus importante pour le responsable de l'évaluation durant le stade de planification. Cette étape a beau sembler évidente, elle est souvent ignorée ou traitée comme peu importante. Il est tout particulièrement important de définir clairement la finalité lorsque l'on évalue des travaux normatifs du fait de la nature délicate du travail et de sa primauté dans le système des Nations Unies. Quelle est la raison d'être de l'évaluation ? Comment les informations de l'évaluation seront-elles utilisées ? L'évaluation est-elle tournée vers l'avenir et axée sur des améliorations, ou bien est-elle destinée à déterminer la valeur d'initiatives normatives



passées seulement ? Les réponses à ces questions façonneront l'approche, la conception et les méthodes de l'évaluation.

45. La plupart des évaluations du travail normatif dans le système des Nations Unies ont une finalité double : *la redevabilité* et *l'apprentissage*, et il y a souvent une tension entre les deux. La conception d'une évaluation pour laquelle l'accent est mis sur la redevabilité est susceptible d'être différente de celle qui met l'accent sur

l'apprentissage et l'amélioration. La conception en vue de la redevabilité est relativement simple ; le responsable de l'évaluation peut avoir recours aux critères de l'OCDE/du CAD (pertinence, efficacité, efficience, durabilité et impact), formuler des questions autour de chacun de ces critères, et structurer l'évaluation en conséquence. Dans un cas de ce type, la participation des parties prenantes devient secondaire. Toutefois, si l'évaluation est tournée vers l'avenir et axée sur l'apprentissage, il y a une foule de questions qui pourraient être examinées (Encadré 5), dont certaines, mais pas toutes, s'inscriront dans les critères standard de l'OCDE/du CAD¹⁷. Dans ce cas, la participation des utilisateurs de l'évaluation prévus est cruciale.

46. Le responsable de l'évaluation devrait, dans la mesure du possible, parvenir à un équilibre entre les deux finalités. Les bailleurs de fonds qui financent les évaluations dans le système des Nations Unies sont souvent impulsés par des préoccupations liées à la redevabilité, comme les résultats obtenus et le rapport résultats/coûts. Cependant, il est souvent possible de présenter des arguments convaincants pour orienter l'évaluation du travail normatif vers l'apprentissage pour l'amélioration des programmes, au vu de la nature non linéaire et souvent qualitative du travail et le manque d'informations sur les pratiques efficaces. Il ne s'agit pas d'abandonner les préoccupations relatives à la redevabilité mais plutôt de parvenir à un juste milieu.

47. Ici encore, tout dépend des circonstances. Par exemple, si des responsables supérieurs envisageaient d'élargir l'échelle du travail normatif ou de le faire évoluer vers une nouvelle phase, alors la finalité de l'évaluation pourrait être d'éclairer ces décisions en identifiant les conditions requises pour que le travail normatif porte ses fruits à l'avenir. Par exemple, la finalité de l'évaluation à mi-parcours *Better Work Haiti* par l'OIT était « d'examiner les progrès effectués vers la réalisation des objectifs immédiats du projet et d'identifier les enseignements tirés de ses services clés mis en œuvre jusque-là, de son approche concernant le renforcement des capacités des parties prenantes et de sa collaboration générale avec les parties prenantes nationales et internationales » (Nexus Associates, Inc., 2012). D'un autre côté, si le programme touchait à sa fin sans aucune chance d'être reproduit, la finalité première pourrait être de déterminer les résultats obtenus. L'Encadré 4 met en relief d'autres utilisations potentielles pour l'évaluation du travail normatif au sein du système des Nations Unies.

« Bien que la plupart des décideurs et des parties prenantes s'intéressent avant tout à l'impact obtenu au niveau des bénéficiaires... les interventions de politique générale sont principalement conçues pour engendrer des changements durables au niveau institutionnel (gouvernement) (effet de « deuxième ordre »), avec seulement un effet indirect au niveau des bénéficiaires. »

- Leeuw and Vaessen, 2009, p. 9

¹⁷ <www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>

Encadré 4. Utilisations des évaluations du travail normatif de l'ONU

- Identifier les enseignements sur les performances afin d'améliorer le programme
- Prendre des décisions sur l'élargissement de l'échelle ou la reproduction du travail normatif et les ressources requises
- Mieux comprendre les forces qui impulsent l'adoption et la mise en œuvre ou la non-adoption et la non-mise en œuvre de la ou des norme(s)
- Faciliter la prise de décisions organisationnelles, comme sur la décentralisation de la responsabilité du travail normatif
- Déterminer les partenariats et les calendriers requis pour parvenir aux effets et à l'impact normatifs souhaités

48. Une fois que la finalité de l'évaluation est claire, les questions auxquelles celle-ci doit s'efforcer de répondre suivront. Ici, le responsable de l'évaluation doit être discipliné et demander : « Quelles sont les *trois* questions les plus importantes auxquelles cette évaluation peut et devrait répondre ? ». En fin de compte, il est possible qu'il y en ait plusieurs autres, mais les trois premières devraient être sans équivoque. Il devrait s'agir des questions les plus étroitement liées à la finalité de l'évaluation, et il devrait être clair pour quiconque lirait le cahier des charges que ces trois questions constituent la priorité et que l'évaluation ne revêtira une valeur que si elle y répond.
49. Quelles sont ces questions clés ? Cela dépend des circonstances. C'est ici qu'un groupe de référence ou une mission de lancement peuvent se révéler utiles. On pourrait être tenté par l'option de se tourner en premier lieu vers les critères de l'OCDE/du CAD pour les questions. Le risque est que l'évaluation comporte au final trop de questions et beaucoup d'informations à recueillir qui ne présentent guère d'utilité sur le plan pratique, et pas assez d'informations sur les aspects les plus importants ; en bref, une évaluation riche en données mais pauvre en information.
50. Le Tableau 5 présente des exemples de questions pertinentes pour les trois catégories générales de travaux normatifs entrepris dans le cadre du système des Nations Unies.

Tableau 5. Questions d'évaluation clés par catégorie de travail normatif

Catégories	Exemples de questions clés
1. Soutien à l'élaboration d'un standard international	<ul style="list-style-type: none"> • Quel(le) a été la stratégie/le processus pour élaborer le standard, et quel succès a-t-il/elle remporté ? • Quels résultats ont été obtenus sur le plan du renforcement des capacités des États membres à prendre part à des forums internationaux pour élaborer le standard ? Quel a été le degré d'inclusivité de la participation ? • Quel a été le degré de succès de l'intervention au moment de parvenir à un consensus sur le standard avec les États membres s'appropriant le processus et les produits ? • Dans quelle mesure le soutien de l'organisation des Nations Unies concernée a-t-il été inclusif ?
2. Soutien pour encourager les gouvernements à intégrer le standard à la législation, aux politiques et aux plans de développement	<ul style="list-style-type: none"> • Dans quelle mesure les interventions se sont-elles révélées efficaces au moment de convaincre les gouvernements d'intégrer le standard international à la législation, aux politiques générales et aux plans de développement ? • Quelles sont les stratégies, les méthodes et les approches qui ont bien fonctionné, et celles qui ont moins bien fonctionné, et pourquoi ? • Dans quelle mesure l'intervention a-t-elle réussi à identifier et à aider les États membres à gérer les risques et à surmonter les barrières à l'intégration du standard dans la législation, les politiques et les plans ?
3. Soutien pour aider les gouvernements à mettre en œuvre la législation, les politiques générales et les plans de développement sur la base du standard international	<ul style="list-style-type: none"> • Dans quelle mesure l'intervention a-t-elle été efficace au moment de développer les capacités des États membres à préparer et à mettre en œuvre la législation, les politiques générales et les plans de développement ayant trait au standard ? • L'organisation de l'ONU avait-elle les bons partenaires et le bon mélange de connaissances techniques spécialisées pour apporter un soutien technique ? • Dans quelle mesure les effets ont-ils été durables ? • Quel a été l'impact au niveau de la population ?

51. Parmi les questions clés liées à un travail normatif faisant partie d'une initiative de plus grande envergure peuvent figurer les suivantes :

- Dans quelle mesure le personnel des Nations Unies était-il conscient de la fonction normative de l'organisation et a-t-il dû rendre des comptes sur l'obtention de résultats normatifs ?
- Dans quelle mesure le rôle normatif a-t-il été bien intégré dans la planification de l'intervention, et également en termes de ressources allouées et de suivi régulier ?
- Quels sont les résultats qui ont été établis et obtenus de manière spécifique par rapport au travail normatif ?
- Dans quelle mesure le travail normatif a-t-il contribué à l'efficacité et à l'efficacé de l'intervention ou des interventions à laquelle ou auxquelles il est lié ?

52. Un exemple concret de ce cas de figure peut être tiré de la Quatrième étude globale des performances du Fonds pour l'environnement mondial (FEM) en 2010. Parmi ses questions clés figuraient :
- Dans quelle mesure le FEM a-t-il suivi les lignes directrices des conventions [Comme la Convention sur la diversité biologique et la Convention sur les polluants organiques persistants] pour lesquelles il est un instrument financier ?
 - Dans quelle mesure le FEM a-t-il été en mesure de fournir un feedback aux conventions sur leurs conseils, la mise en œuvre de ces conseils et les résultats obtenus, notamment sur les questions concernant des domaines multiples ?
 - Le système de gouvernance du FEM est-il adéquat et conforme aux standards internationaux ?
53. Parmi les questions visant à évaluer le travail de plaidoyer mené pour convaincre les gouvernements d'adopter un standard international pourraient figurer les suivantes :
- Dans quelle mesure les initiatives de plaidoyer de l'intervention se sont-elles fondées sur la recherche-action, le feedback reçu des groupes cibles et d'autres pratiques de plaidoyer efficaces connues ?
 - Dans quelle mesure la stratégie de plaidoyer/communication s'est-elle révélée efficace au moment de formuler la question et de cibler les décideurs clés ?
 - Dans quelle mesure l'agence de l'ONU a-t-elle adapté ses activités de plaidoyer de manière opportune et continue pour répondre à l'environnement politique et économique en mutation ?
54. Si l'évaluation évalue des produits normatifs, des bases de données et des ressources associées, alors il devrait y avoir des questions sur la qualité, l'accès et la distribution (voir, par exemple, les critères d'évaluation dans l'annexe 8 de *l'Évaluation des activités de la FAO en matière de renforcement des capacités en Afrique* [FAO, 2010]).
55. Le guide de l'UNEG sur l'intégration des droits de l'homme et de l'égalité des sexes dans l'évaluation stipule que « toutes les interventions de l'ONU ont pour mandat d'aborder les questions liées aux DH [droits de l'homme] et à l'EGS [égalité des sexes] » (UNEG, 2011, p. 16). Il s'ensuit que les évaluations des travaux normatifs de l'ONU devraient englober une appréciation des dimensions des interventions ayant trait aux droits de l'homme et au genre. Voici deux exemples de questions pour cette finalité :
- Dans quelle mesure l'intervention a-t-elle incorporé des principes relatifs aux droits de l'homme et a-t-elle intégré des considérations d'égalité des sexes ?
 - Dans quelle mesure les partenaires de la mise en œuvre ont-ils eu des comptes à rendre sur l'incorporation des droits de l'homme et de l'égalité des sexes dans leurs travaux normatifs ?
56. Les questions relatives à la conception de l'intervention sont souvent importantes parce que la mise en œuvre insuffisante est souvent liée à des défauts dans la conception. Il en va de même pour les partenariats intervenant dans les travaux normatifs. L'intervention avait-elle les bons partenaires, dotés des connaissances techniques optimales ? Quel a été le degré d'appropriation et d'adoption parmi les partenaires ? L'identification et l'atténuation des

risques constituent une partie importante de la conception, de sorte qu'une évaluation du travail normatif pourrait englober des questions relevant de ce domaine.

57. N'oubliez pas que la plupart des évaluations disposent de ressources limitées : rares sont celles qui peuvent se permettre d'évaluer de nombreuses questions complexes dans le cadre d'une seule évaluation. Mieux vaut répondre à quelques questions clés de façon complète qu'en couvrir un grand nombre de façon superficielle. Le mot d'ordre devrait être la qualité, et non la quantité. La responsabilité de concentrer l'évaluation sur ce qui est le plus important et réalisable, au vu du temps et des ressources disponibles, incombe principalement au responsable de l'évaluation.

Tâche 2.3 : Sélectionner les critères autour desquels bâtir l'évaluation

58. Le système des Nations Unies utilise les critères d'évaluation de l'OCDE/du CAD, à savoir : pertinence, efficacité, efficience, impact et durabilité, pour bâtir ses évaluations, comme l'illustre l'Encadré 5.

Encadré 5. Critères d'évaluation et questions normatives de l'OCDE/du CAD

Pertinence

- Le travail normatif a-t-il satisfait les besoins que ses bénéficiaires/les garants d'obligations ont exprimés et dont ils sont convenus ?
- Le travail normatif a-t-il un sens en ce qui concerne les conditions, les besoins ou les problèmes auxquels il est censé remédier ?
- Le travail normatif s'inscrit-il dans le mandat et les priorités mondiales de l'organisation de l'ONU ?
- Le travail normatif s'aligne-t-il sur les priorités et les politiques générales du gouvernement du pays d'accueil ?
- Le travail normatif s'aligne-t-il sur le Plan cadre des Nations Unies pour l'aide au développement (UNDAF) ?

Efficacité

- Quels effets directs le travail normatif a-t-il obtenus, prévus et imprévus, positifs et négatifs ?
- Le travail normatif touche-t-il les bénéficiaires, les titulaires de droits et les garants d'obligations prévus ?
- Quels seraient les effets probables si l'intervention n'avait pas eu lieu ?
- Le travail normatif s'est-il suffisamment reflété dans le modèle logique, la théorie du changement et le cadre de mesure des performances du programme ?
- Le travail normatif obtiendra-t-il vraisemblablement les effets une fois l'initiative terminée ?
- Quel a été le degré d'efficacité du processus de l'organisation de l'ONU en ce qui concerne la participation des parties prenantes, l'inclusivité, l'accessibilité, la transparence et le développement des capacités ?
- Dans quelle mesure le travail normatif intègre-t-il l'égalité des sexes, l'autonomisation des femmes, les droits de l'homme et la coopération Sud-Sud ?
- Quelle a été la plus-value apportée par l'organisation ?
- Dans quelle mesure les produits normatifs distribués par l'organisation de l'ONU, y compris les produits de connaissances et les ressources d'information, sont-ils utilisés par les diverses parties prenantes ?

Efficiences

- Les ressources et les intrants (fonds, connaissances techniques spécialisées et temps) sont-ils utilisés pour obtenir les résultats normatifs de manière efficiente et opportune ?
- Le rapport entre les coûts et les résultats est-il raisonnable ?
- Y a-t-il des ressources suffisantes pour parvenir aux effets prévus du travail normatif ?
- Les coûts sont-ils similaires à ceux de travaux normatifs analogues ?
- Les variances entre les dépenses planifiées et réelles sont-elles justifiées ?
- Quel est le degré d'efficacité des mécanismes utilisés pour le dialogue et la coopération intergouvernementaux ?

Durabilité

- Quelle est la probabilité que les avantages du travail normatif continueront d'exister longtemps après son achèvement et sans sursolliciter les organisations locales et les organisations partenaires ?
- Les capacités en cours de développement sont-elles adéquates pour s'assurer que les institutions/organisations prendront la relève et maintiendront les avantages envisagés ?
- Dans quelle mesure le programme présente-t-il une résilience inhérente face aux risques futurs ?
- Dans quelle mesure les partenaires et les bénéficiaires participent-ils aux résultats normatifs et se les « approprient »-ils ?
- Le travail normatif adhère-t-il à des pratiques et standards environnementaux durables ?

Impact

- Dans quelle mesure le travail normatif a-t-il contribué à l'obtention de résultats au niveau de l'impact ?
- Quel impact négatif le travail normatif a-t-il produit, directement ou indirectement, prévu ou imprévu ?

59. Les critères et les questions du CAD/OCDE devraient servir de guide. Si l'évaluation doit inclure tous les critères, le cahier des charges devra clairement stipuler ceux qui sont prioritaires. Une façon d'éviter de trop exiger de l'évaluation consiste à couvrir certains des critères de l'OCDE/du CAD – la pertinence et l'efficacité, par exemple – au moyen d'une étude documentaire et éventuellement d'un entretien avec un groupe de réflexion, ce qui laisse plus de ressources permettant de traiter en profondeur les critères restants.
60. Au moment de bâtir l'évaluation d'un travail normatif, le responsable de l'évaluation peut appliquer d'autres critères. Certains pourraient être d'ordre technique, liés par exemple aux standards scientifiques (p. ex. qualité scientifique (*quality of science*)¹⁸ sous-jacente aux normes d'établissement des standards) ; d'autres pourraient avoir trait aux droits de l'homme, à l'autonomisation des femmes et aux principes relatifs au comportement dans les interventions humanitaires, par exemple. Certaines évaluations pourraient choisir de se concentrer sur les questions de gestion, comme la mesure dans laquelle la décentralisation influe sur le travail normatif, ou sur les capacités techniques et managériales dont ont besoin les gouvernements pour mettre en œuvre les standards. L'évaluation pourrait être adaptée pour répondre aux éléments les plus importants. L'Encadré 6 donne un bon exemple d'un ensemble compact de questions d'une évaluation indépendante du Programme de production plus propre (PP) ONUDI-PNUE qui vise à aider des institutions d'accueil nationales à fournir des services de PP conformes aux standards internationaux reconnus.

¹⁸ *Quality of science* n'a pas de définition internationale standard. Parmi les aspects de la qualité mis en relief dans les écrits on peut citer : (i) des résultats qui parviennent aux objectifs et une documentation adéquate des méthodes et des résultats (Commission de réglementation nucléaire des États-Unis, 2005) ; et (ii) reproductibilité, caractère nouveau, indépendance, méthodes et clarté de la contribution (Gill et Gilbert, 2011).

Encadré 6. Questions pour bâtir l'évaluation d'un programme conçu pour renforcer le soutien institutionnel national apporté aux standards de production plus propre

Pertinence : Les éléments du programme (c.-à-d. le concept de production plus propre (PP), les services de PP, les Centres nationaux des technologies de production plus propre, le réseau mondial et les intrants en termes d'aide technique) sont-ils applicables et utiles pour les bénéficiaires prévus (c.-à-d. le secteur privé, le gouvernement, le milieu universitaire et les instituts de recherche dans le pays d'accueil) ?

Efficacité : La conception du programme (c.-à-d. centres nationaux, gestion et travail en réseaux à l'échelle mondiale et assistance technique) et sa mise en œuvre permettent-elles aux Centres et aux bénéficiaires d'obtenir les résultats prévus du programme (c.-à-d. adoption de la PP) ?

Efficience : Le programme est-il conçu et mis en œuvre de manière à obtenir les avantages optimaux à partir des ressources dont il dispose ? Les Centres et les autres activités de programme potentielles sont-ils gérés et administrés de manière à favoriser la prestation de services aux bénéficiaires ?

Durabilité : Est-il probable ou vraisemblable que les avantages (p. ex. disponibilité des services de PP, bénéfices environnementaux et sur le plan de la productivité dans l'industrie, etc.) obtenus grâce au programme se poursuivront à l'avenir ?

Développement des capacités : Le programme développe-t-il les capacités essentielles (p. ex. en ce qui concerne la productivité des ressources, la gestion environnementale, l'entrepreneuriat et/ou les partenariats publics/privés) pour que les principales parties prenantes puissent améliorer leur bien-être actuel et futur ?

Appropriation : Les parties prenantes locales considèrent-elles que le programme leur appartient et s'engagent-elles à faire avancer les buts et objectifs du programme et à agir en fonction de ses produits ?

Source : Organisation des Nations Unies pour le développement industriel et Programme des Nations Unies pour l'environnement (2008). *Évaluation indépendante ONUDI-PNUE du Programme de production plus propre*. Vienne : ONUDI, pp. 10-11.

Tâche 2.4 : Préparer les indicateurs

61. L'idéal est que le responsable de l'évaluation ait une théorie du changement qui saisisse les effets et l'impact prévu du travail normatif, ainsi que les suppositions sous-jacentes. Il peut y avoir un cadre de mesure des performances ou une analyse basée sur un cadre logique avec des indicateurs. Mieux encore, l'intervention peut avoir quelques données utiles relatives au suivi des performances basées sur les indicateurs. Cependant, en réalité, il est tout à fait possible qu'aucun de ces éléments n'existe, ou si certains sont présents, qu'ils présentent de sérieux défauts, auquel cas le responsable de l'évaluation devra peut-être préparer une théorie du changement rétrospective et un ensemble nouveau ou revu d'indicateurs pour chaque énoncé de résultat et d'impact. Il est supposé que la plupart des évaluations de travaux normatifs dans le système des Nations Unies tentent de transmettre un message sur l'efficacité. Il s'agit en général d'évaluer la mesure dans laquelle l'intervention a obtenu ses résultats prévus, d'où le besoin d'indicateurs. Les indicateurs doivent répondre à la question suivante : « Qu'attendrions-nous de voir comme preuve vérifiable de l'obtention de cet effet direct ou de cet impact ? »

62. L'élaboration d'indicateurs pour le travail normatif n'est pas tâche aisée. Souvenez-vous du temps et des efforts consacrés à la mise au point des indicateurs pour les OMD. Quelques règles générales s'appliquent :

- **Limitez le nombre d'indicateurs à deux ou trois par énoncé d'effet direct/d'impact.** Ainsi, le volume de données à recueillir sera maintenu à un niveau gérable. Il ne faut pas laisser la collecte de données devenir trop laborieuse ou coûteuse. Si un ou deux bons indicateurs peuvent fournir des preuves concluantes du changement prévu, n'allez pas plus loin.
- **Incluez au moins un indicateur qualitatif par énoncé de résultat.** Les indicateurs qualitatifs permettent à l'évaluateur de mieux exploiter les informations et d'en extraire des nuances importantes.
- **Veillez à ce que les indicateurs soient sensibles au genre,** dans la mesure du possible. Un indicateur sensible au genre est une mesure du changement au fil du temps par rapport à l'égalité des sexes, et une mesure qui produit des données ventilées par sexe. Par exemple : *Mesure dans laquelle les points de vue des femmes et des hommes sont entendus dans les sessions de dialogue sur les politiques générales.*
- **Pensez SMART:** Les indicateurs devraient être spécifiques (*specific*), mesurables (*measurable*), atteignables (*attainable*), pertinents (*relevant*) et, si possible, limités dans le temps (*time-bound*).

63. Le Tableau 6 ci-dessous propose des exemples d'indicateurs pertinents pour une campagne de plaidoyer soutenue par une organisation de l'ONU et visant à encourager les États membres à adopter une norme internationale.

Tableau 6. Indicateurs pour des activités de plaidoyer de nature normative

Effets	Indicateurs
1. Soutien renforcé pour la norme internationale	<ul style="list-style-type: none"> • Variation du pourcentage de personnes conscientes de la question normative de la campagne parmi les groupes cibles ; nombre de groupes cibles (p. ex. décideurs et leaders d'opinion) conscients des principes de la campagne et des messages clés liés à la norme • Nombre de groupes de parties prenantes clés d'accord sur la définition du problème lié à la norme • Mesure de la couverture médiatique de la question normative • Diversité des groupes intervenant sur la question telle que mesurée par le nombre d'« alliés improbables » qui soutiennent la norme • Nombre de <i>followers</i> sur Twitter, de visiteurs discrets du site Web de la campagne, de mentions mensuelles dans les médias
2. Capacités organisationnelles des partenaires qui soutiennent la norme renforcées	<ul style="list-style-type: none"> • Nombre d'améliorations en matière de gestion parmi les organisations partenaires pour soutenir la mise en œuvre de la norme • Aptitude des organisations partenaires à mettre en œuvre la norme • Degré de stabilité des organisations partenaires pour mettre en œuvre la norme

3. Alliances qui soutiennent la norme renforcées	<ul style="list-style-type: none"> • Nombre d'alliances stratégiques qui soutiennent la norme • Degré de collaboration parmi les partenaires pour soutenir la norme • Envergure et forces des coalitions formées pour soutenir la norme
4. Base de soutien renforcée pour la norme	<ul style="list-style-type: none"> • Diversité des groupes intervenant sur la question telle que mesurée par le nombre d'« alliés improbables » qui soutiennent la norme • Mesure de la couverture médiatique de la question normative • Nombre de groupes cibles (p. ex. décideurs et leaders d'opinion) conscients des principes de la campagne et des messages clés liés à la norme • Nombre de <i>followers</i> sur Twitter, de visiteurs discrets du site Web de la campagne, de mentions mensuelles dans les médias
5. Politiques qui reflètent la norme améliorées ¹⁹	<ul style="list-style-type: none"> • Données documentées qui montrent qu'une nouvelle politique sur la question de la campagne est en cours de mise au point • Une nouvelle politique qui traduit la norme est adoptée sous forme d'ordonnance, de législation, d'accord contraignant • Niveau de financement appliqué à la mise en œuvre de la politique nouvelle ou améliorée

Adapté de Reisman, Gienapp et Stachowiak, 2007.

64. La collaboration avec les parties prenantes peut aider à surmonter la difficulté de préparer les indicateurs pour les résultats normatifs. Dans l'idéal, le responsable de l'évaluation devrait former une équipe à cette fin et l'équipe devrait englober des spécialistes du thème et au moins un manieur de mots doté d'une bonne expérience en matière d'indicateurs. C'est une série de courts ateliers de travail qui convient le mieux pour cette tâche parce qu'il est probable que les indicateurs devront passer par plusieurs moutures avant d'être prêts. L'idéal est que les parties prenantes clés, y compris les responsables dont le travail normatif est en cours d'évaluation, valident les indicateurs une fois qu'ils ont été rédigés.
65. Lors de la préparation des indicateurs, n'ignorez pas les aspects qui peuvent sembler évidents. Par exemple, si l'un des effets est « Adoption accrue de normes relatives à l'innocuité des aliments par les États membres », un indicateur simple pourrait être « nombre de pays ciblés qui ont adopté les standards durant la période examinée ». Cependant, cela est insuffisant pour présenter un compte rendu des performances, car il peut y avoir eu quelques États membres qui ont adopté le standard pour des raisons autres que l'intervention. Dans ce cas, un autre indicateur est requis pour que l'évaluateur ait une deuxième source de données factuelles. Par exemple : « Nombre d'experts et d'ONG partenaires qui pensent que l'adoption des standards n'aurait pas eu lieu sans l'intervention ». Les évaluateurs devraient toujours chercher des mesures simples, certes, mais parfois ce sont des mesures complexes

¹⁹ NB : Il s'agit là d'un résultat à moyen terme (intermédiaire), voire à long terme.

À travers des entretiens semi-structurés dans plusieurs pays africains en 2009, les évaluateurs chargés d'estimer le développement des capacités de la FAO ont constaté que nombre de parties prenantes clés n'étaient pas conscientes de certains des produits normatifs clés de la FAO et/ou éprouvaient des difficultés à y accéder via Internet. Cette constatation a été une surprise. L'accès aux produits normatifs et leur distribution sont devenus des indicateurs de substitution utiles des performances normatives et ont abouti à des recommandations pratiques visant à remédier aux insuffisances.

FAO, 2010

qui sont nécessaires pour présenter un compte rendu des performances, particulièrement lors de l'évaluation de l'impact de travaux normatifs.

66. N'oubliez pas que les énoncés des effets et les indicateurs devraient se concentrer sur l'essentiel du travail normatif : ce qui est contrôlé par l'organisation chargée de la mise en œuvre et qui peut être raisonnablement prédit. Étant donné que la plupart des travaux normatifs ont lieu dans des environnements dynamiques, il y aura probablement des effets imprévus et de nouveaux indicateurs qui seront révélés durant l'évaluation.

67. La plupart des rapports de lancement d'une évaluation comportent une matrice de l'évaluation qui présente les questions clés, les sous-questions et

les indicateurs correspondants, les sources des données et les méthodes d'évaluation. Le Tableau 7 donne un exemple de la section portant sur l'efficacité d'une matrice de l'évaluation pour un programme de promotion de standards relatifs à la biodiversité.

Tableau 7. Matrice d'évaluation des résultats normatifs liés aux standards en matière de biodiversité

Critère d'efficacité	Question clé et sous-questions	Indicateurs	Sources des données	Méthodes
<p><u>Effets :</u></p> <p>1. Adoption accrue des standards en matière de biodiversité par les gouvernements nationaux participants.</p> <p>2. Prise de conscience accrue des standards en matière de biodiversité par les responsables nationaux de l'environnement.</p>	<p>Quels résultats l'intervention a-t-elle obtenus ?</p> <ul style="list-style-type: none"> ▪ Quels seraient les effets probables si l'intervention n'avait pas eu lieu ? ▪ Quels résultats imprévus l'intervention a-t-elle obtenus ? • Quels sont les facteurs qui améliorent et entravent les résultats ? 	<p>1.1 Nombre de gouvernements qui ont adopté les standards</p> <p>1.2 Nombre de gouvernements qui ont promulgué des lois ou des réglementations reflétant les standards</p> <p>1.3 Perceptions des partenaires, y compris les ONG environnementales, concernant la contribution de l'organisation de l'ONU à l'adoption accrue et aux facteurs qui améliorent et entravent l'adoption.</p> <p>2.1 Nombre de responsables (de sexe masculin/de sexe féminin) qui ont assisté à des événements de sensibilisation et qui signalent une sensibilisation accrue</p> <p>2.2 Perceptions d'autres parties prenantes, y compris les ONG environnementales, concernant la prise de conscience des standards par les responsables environnementaux</p>	<p>Documents du programme</p> <p>Autorités nationales</p> <p>Responsable et personnel du programme</p> <p>Représentants d'ONG</p> <p>Experts et groupes qui n'ont pas pris part à l'intervention</p>	<p>Examen des documents</p> <p>Entretiens</p> <p>Discussions en groupes de réflexion</p>

68. Si l'intervention a préparé un cadre de mesure des performances, il est possible qu'il y ait des données de référence disponibles, ce qui se révélerait utile pour les évaluateurs ; cela les aiderait en effet à évaluer les changements qui ont eu lieu depuis le commencement de l'intervention. Lorsque l'organisation de l'ONU cherche à établir un standard, et qu'aucun standard n'était en place précédemment, la base de référence sera zéro. Lorsqu'il n'y a pas de données de référence fiables, les évaluateurs doivent tenter de reconstruire la base de référence, en général en procédant à un examen des documents et en organisant des entretiens avec les responsables. Cela peut être une entreprise difficile, tout particulièrement délicate si les archives sont peu précises, si les personnes qui étaient là au début de l'intervention sont passées à autre chose et si le contexte de la programmation est dynamique.

Tâche 2.5 : Sélectionner la conception et la méthodologie appropriées pour l'évaluation

69. Le responsable de l'évaluation dispose de nombreux choix pour concevoir l'évaluation d'un travail normatif et pour sélectionner les méthodologies. En premier lieu, le responsable de l'évaluation doit décider du moment auquel mener l'évaluation, parce que le *timing* influence le type d'évaluation à mener. Il s'agit là d'une question importante parce que les résultats normatifs peuvent être à long terme, émergents et imprévisibles, comme dans le cas d'initiatives visant à élaborer une norme internationale, et d'autres initiatives qui aident les États membres à aligner leurs politiques et leur législation sur les conventions internationales. Il n'y a pas de référence absolue : le choix est situationnel et déterminé par la finalité de l'évaluation, ses questions clés, le temps et les ressources disponibles, et la manière dont les données de l'évaluation seront utilisées. Le Tableau 8 donne quelques conseils pour aider les responsables de l'évaluation à effectuer des choix prudents.

Tableau 8. Choix des types et conceptions pour les évaluations de travaux normatifs

Types	Description	Quand l'utiliser
Évaluation sommative	Une évaluation menée à la fin d'une intervention, en général pour fournir des informations sur sa valeur et son impact.	À la fin d'une intervention normative, mais le mieux est de la planifier au commencement de l'intervention.
Évaluation formative	Une évaluation menée vers le milieu d'une intervention, se concentrant généralement sur des questions opérationnelles et visant à identifier les points forts et les défauts. Parfois appelée évaluation de processus ou évaluation à mi-parcours.	Vers le milieu d'une intervention normative, de manière à ce que des actions correctives puissent être prises pour renforcer les effets et l'impact. Le mieux est de la planifier au commencement de l'intervention.
Évaluation développementale ²⁰	Une évaluation qui aide à mettre au point une intervention et à en renforcer l'efficacité durant sa mise en œuvre.	Bien adaptée pour le travail normatif dans des systèmes complexes où les interventions sont expérimentales et les effets peu clairs.
Conceptions		
Expérimentale	Une conception <i>ex-ante</i> demandant l'assignation aléatoire d'une population à aux moins deux groupes, dont l'un qui servira de groupe témoin ou contrefactuel.	À utiliser chaque fois que c'est possible et financièrement abordable afin de déterminer si l'impact peut être assigné à l'intervention comme cause unique, c.-à-d. si le travail normatif était à la fois nécessaire et suffisant pour engendrer l'impact ²¹ .
Quasi-expérimentale	Une conception dans le cadre de laquelle un groupe d'intervention et un groupe de non-intervention, ou témoin (contrefactuel), sont formés soit avant soit après l'intervention, mais sans assigner aléatoirement la population à un	À utiliser lorsque ce degré de rigueur est nécessaire et faisable pour prouver que le travail normatif était à la fois nécessaire et suffisant pour entraîner un changement net au niveau de l'impact. Difficile de trouver des groupes de

²⁰ Voir le Chapitre 1 pour un complément d'informations sur l'évaluation développementale et un exemple de situation dans laquelle ce type d'évaluation est approprié.

²¹ Seulement une (Coulibaly, 2010) des plus de 30 évaluations de travail normatif examinées pour ce manuel a utilisé une conception expérimentale ou quasi-expérimentale.

	groupe.	comparaison qui soient bien assortis.
Non-expérimentale, y compris approches basées sur la théorie et les cas concrets.	Une conception de l'évaluation dans laquelle on ne cherche pas à comparer des groupes d'intervention et témoins, et où l'accent est mis sur la description.	À utiliser à un moment unique pour évaluer ce qui était planifié, ce qui a eu lieu, ce qui a changé et pourquoi.

70. Bien que les conceptions expérimentales et quasi-expérimentales puissent constituer des moyens puissants de répondre aux questions relatives à l'efficacité, et en particulier à celles ayant trait à l'impact, elles sont particulièrement adaptées aux conditions dans lesquelles il y a une cause primaire et un effet principal. Une grande partie des travaux normatifs de l'ONU, toutefois, font intervenir de nombreux acteurs, de nombreuses causes potentielles et tout autant d'effets possibles. Pensez, par exemple, aux initiatives de l'ONU visant à convaincre les gouvernements d'États d'accepter les lois internationales sur le travail, les conventions relatives aux droits de l'homme ou les protocoles de développement durable²². Un « bloc causal » intervient généralement dans les changements normatifs de ce type ; le contenu de ce bloc est susceptible d'être différent d'un pays à l'autre ; et le « succès » peut prendre des formes différentes selon le pays.

« La plupart des éléments de l'évaluation ont utilisé de petits échantillons ciblés de cas riches en informations, au lieu de gros échantillons aléatoires qui permettent des généralisations valides et des essais significatifs sur le plan statistique. »

- *Évaluation de la stratégie de plaidoyer et de mobilisation sociale UNICEF-gouvernement du Pakistan* (Stiles, 2002, p. 9)

71. Il est tout particulièrement difficile de déterminer le « succès » dans les initiatives conçues pour élaborer une norme ou un standard. Il faut parfois faire des compromis pour parvenir au consensus, et toutes les normes ou tous les standards ne sont pas réalisables pour les pays en développement. Lorsque le tableau n'est pas entièrement noir et blanc, il faut mener une enquête qualitative avec des questions de causalité visant à rendre les diverses nuances de gris ; par exemple : « Qu'est-ce qui fonctionne pour qui, pourquoi et dans quelles conditions ? » (Pawson et Tilley, 1997).

72. Quelles méthodes le responsable de l'évaluation devrait-il utiliser pour évaluer le travail normatif ? Il n'y a pas de méthode unique idéale ou de combinaison unique idéale de méthodes : tout dépend de la finalité de l'évaluation, de son axe, du temps et des ressources disponibles et de l'utilisation prévue des informations de l'évaluation. Ceci dit, dans la plupart des cas, quelques **règles générales** s'appliquent :

1. Utilisez des **méthodes multiples** afin de compenser les faiblesses de toute méthode unique. Un mélange de méthodes confère une dimension de complémentarité à

²² Pour un examen approfondi des conditions strictes à satisfaire pour les conceptions expérimentales et quasi-expérimentales, voir Stern *et al.* (2012). « Broadening the Range of Designs and Methods for Impact Evaluations ». *Department for International Development Working Paper No. 38*. Londres : Department for International Development.

- l'évaluation : les résultats obtenus d'une méthode peuvent souvent aider l'évaluateur à mieux comprendre les résultats obtenus grâce à une autre.
2. Dans la plupart des circonstances, il faut des **données qualitatives ainsi que quantitatives** pour évaluer le travail normatif afin de révéler les différents éléments du bloc causal et d'examiner les questions qualitatives de manière plus approfondie.
 3. **Des champs d'enquête et des sources de données multiples** sont conseillés parce qu'ils permettent la **triangulation**, « l'utilisation d'au moins trois théories, sources ou types d'informations, ou démarches d'analyse, pour vérifier et soutenir une appréciation », renforçant ainsi la confiance dans les constatations. (CAD/OCDE, 2002, p. 37).²³
73. Suivant les règles générales ci-dessus, les Tableaux 9, 10, 11 et 12 évaluent les méthodes pour chaque catégorie de travail normatif, en commençant par les initiatives qui soutiennent l'élaboration d'une norme, d'un standard ou d'une convention international(e).

Tableau 9. Méthodes d'évaluation pour les initiatives qui soutiennent l'élaboration d'une norme, d'un standard ou d'une convention international(e)

Méthodologie	Avantages	Inconvénients
1. <i>Analyse documentaire</i> axée sur des types similaires de soutien et d'évaluations. Devrait englober certaines ressources sélectionnées sans la liste située à la fin chaque chapitre de ce manuel.	Pourrait révéler des enseignements concernant les pratiques efficaces et inefficaces dans des circonstances similaires.	Coûteuse sauf si elle est concentrée et/ou entreprise par un évaluateur subalterne.
2. <i>Examen des documents</i>	Généralement nécessaire pour comprendre le contexte, les parties prenantes, les rôles, les activités et les dépenses. Utile au moment de répondre aux questions ayant trait à la pertinence et à l'efficacité.	Demande beaucoup de temps. Peut ne pas engendrer beaucoup de données liées aux indicateurs des effets et de l'impact.

²³ Par exemple, une évaluation de 2007 du travail en matière d'affaires politiques mené par le Conseil de sécurité des Nations Unies a eu recours à des enquêtes auto-administrées, à des entretiens en personne, à des analyses des données des programmes, à des études de cas et à une analyse documentaire, s'inspirant de 11 sources de données différentes (Conseil économique et social des Nations Unies, 2007)

<p>3. <i>Entretiens avec des interlocuteurs (sources d'information) clés</i></p>	<p>Le moyen le plus efficace de recueillir des informations qualitatives approfondies, p. ex. en examinant ce qui a donné de bons résultats, ce qui n'a pas fonctionné et pourquoi.</p>	<p>Demandent beaucoup de temps. Peuvent être coûteux si les interlocuteurs clés sont nombreux et éloignés. Les entretiens téléphoniques et sur Skype sont généralement moins efficaces que les entretiens face à face mais beaucoup moins coûteux. Difficile de contrôler le biais éventuel. Requièrent des évaluateurs qualifiés et expérimentés et il faut un temps considérable pour organiser les entretiens.</p>
<p>4. <i>Discussion en groupe de réflexion</i></p>	<p>Peut donner lieu à des données qualitatives approfondies. Peut être menée avec différents groupes de parties prenantes pour comparer les points de vue sur une question normative. Une bonne méthode pour obtenir les points de vue de nombreuses personnes en peu de temps. On obtient souvent des résultats tout particulièrement bons lorsque les membres des groupes prennent part à des travaux similaires et/ou au même niveau.</p>	<p>Requiert généralement entre 6 et 10 participants avec des antécédents similaires pendant plusieurs heures, ce qui peut être difficile à organiser. Il faut deux évaluateurs : un pour faciliter, l'autre pour prendre des notes. Doit être bien facilitée pour obtenir la pleine participation. Demande que les membres du groupe soient au même niveau et/ou prennent part à des travaux similaires. Inadapté s'il y a une hiérarchie rigide dans le groupe.</p>
<p>5. <i>Enquête auto-administrée</i></p>	<p>Les enquêtes auto-administrées prennent de nombreuses formes, y compris des enquêtes menées par courrier, par e-mail et en ligne. Les enquêtes par e-mail sont adaptées pour obtenir des informations auprès d'un grand nombre de personnes. Les enquêtes en ligne sont rapides, peu coûteuses et ont une ample portée potentielle. Des logiciels facilement accessibles facilitent le traitement des chiffres pour déterminer les variances, et pour créer des tableaux et des graphiques.</p>	<p>Pour obtenir un bon taux de réponse à partir des enquêtes auto-administrées, les récipiendaires doivent être motivés ou obligés à répondre. Pour les enquêtes en ligne, les taux de réponse chutent souvent après 10 minutes, donc ces enquêtes doivent rester courtes et se limiter principalement à des questions fermées, sauf si des moyens incitatifs sont mis en œuvre. Les résultats ne sont pas représentatifs et sont soumis à un biais potentiel lorsque les participants se sélectionnent eux-mêmes.</p>
<p>6. <i>Observation directe</i>, p. ex. lorsque les États membres se réunissent pour réfléchir à un standard</p>	<p>Utile pour apprécier le processus utilisé pour élaborer une norme ou un standard, la participation des parties prenantes, la qualité des partenariats et le leadership normatif. Devrait être semi-structurée, utilisant un aide-mémoire ou une matrice de notation pour l'évaluation. Il est utile d'avoir deux évaluateurs qui observent en</p>	<p>Difficile de contrôler le biais potentiel. Requiert des évaluateurs expérimentés.</p>

	parallèle et comparent leurs notes respectives par la suite ²⁴ .	
7. <i>Analyse coûts/avantages et analyse coûts/efficacité</i>	Utiles pour répondre aux questions liées à l'efficacité. Peuvent être utilisées pour comparer le coût d'une intervention à celui d'une autre lorsque les conditions sont similaires.	Il peut être difficile de se procurer des données financières exactes. Il peut se révéler nécessaire d'avoir des connaissances techniques financières pour analyser les données. Les évaluateurs doivent faire preuve de circonspection au moment d'interpréter les résultats de l'analyse, car les avantages du travail normatif sont souvent difficiles à quantifier ²⁵ .
8. <i>Cartographie</i> , p. ex. cartographie des parties prenantes ou du processus utilisé pour parvenir à un accord sur un standard ou une norme.	Aide à identifier des parties prenantes multiples, leurs rôles respectifs et les relations entre eux, et les moteurs/entraves du/au changement. Cette méthode peut être mise en œuvre pour se préparer à une évaluation, et peut aussi être employée durant l'évaluation afin de créer des cartes visuelles des parties prenantes et de leurs relations dans chaque pays participant comme moyen d'identifier la variété de difficultés et d'opportunités rencontrés par les États membres pour obtenir un soutien pour les standards.	Pas particulièrement utile lorsque seulement quelques parties prenantes sont impliquées. Difficile à effectuer lorsqu'il y a beaucoup de parties prenantes et/ou lorsque le travail normatif a fait intervenir un processus compliqué sur une longue période.

74. L'évaluation de l'élaboration d'une norme, d'un standard ou d'une convention international(e) est susceptible de se concentrer sur l'évaluation des processus et/ou des structures de gouvernance utilisés pour parvenir à l'accord ou au consensus sur un standard, une norme ou une convention. L'évaluation serait largement qualitative et requerrait des méthodes robustes d'enquête qualitative comme celles du Tableau 9. L'accord ou le consensus sur un standard, une norme ou une convention devraient être assez faciles à déterminer. Un aspect qui pourrait être un peu plus délicat est celui des effets imprévus, tant

²⁴ Dans une évaluation récente, la Division d'inspection et d'évaluation du Bureau des services de contrôle interne a eu recours à l'observation directe parmi les plusieurs méthodes employées pour examiner le travail normatif effectué par le Département des affaires économiques et sociales (Conseil économique et social des Nations Unies, 2011).

²⁵ Pour une discussion approfondie sur les difficultés rencontrées dans le cadre de l'analyse coûts/avantages, voir *Cost-Benefit Analysis in World Bank Projects* (Banque mondiale, 2010), disponible sur : siteresources.worldbank.org/INTOED/Resources/cba_overview.pdf

positifs que négatifs. Les entretiens et les discussions en groupe de réflexion figureraient parmi les meilleures méthodes pour révéler les effets imprévus, et il serait prudent d'inclure certains groupes de parties prenantes qui pourraient avoir été exclus ou qui ont choisi de ne pas participer au processus normatif.

75. Une difficulté méthodologique fréquemment rencontrée pour l'évaluation dans cette catégorie de travail normatif est liée au fait que les représentants des pays ou des organisations partenaires changent fréquemment. Dans ce cas, le responsable de l'évaluation devrait tenter de trouver les personnes qui sont impliquées depuis assez longtemps pour avoir des points de vue historiques éclairés sur le travail normatif. Cela requerrait un échantillonnage ciblé, lequel est légitime pour une enquête qualitative. Bien que les évaluateurs doivent se montrer prudents et ne pas trop généraliser, ils peuvent extrapoler les données recueillies à partir d'un échantillon ciblé. Les extrapolations sont « des spéculations modestes sur l'applicabilité probable des constatations à d'autres situations dans des conditions similaires mais pas identiques » (Patton, 2008, p. 459). Les études de cas peuvent aussi se révéler appropriées si l'évaluation tente de comparer des cas multiples de travaux normatifs visant à élaborer une norme, un standard ou une convention international(e).

Tableau 10. Méthodes d'évaluation pour les initiatives aidant les gouvernements à intégrer un standard international à la législation, aux politiques générales et aux plans de développement

Méthodologie	Avantages	Inconvénients
1. <i>Toutes celles recensées dans le Tableau 9 pourraient s'appliquer</i> ²⁶		
2. <i>Études de cas</i>	Utiles lorsque l'un des objectifs premiers est de parvenir à bien comprendre le contexte de programmation et les raisons pour lesquelles une intervention normative a réussi ou non. En choisissant 4 ou 5 unités de cas et en effectuant une analyse dans le cadre de chaque unité et sur plusieurs unités, il est possible d'obtenir des informations qualitatives et quantitatives importantes sur les conditions requises pour l'intégration réussie des standards internationaux ou sur le type de soutien requis de la part des Nations Unies pour aider les	Demandent beaucoup de temps et peuvent être coûteuses. Il est risqué de généraliser les constatations sur la base de deux ou trois cas seulement. Doivent se concentrer sur les cas riches en information pour justifier l'effort et le coût.

²⁶ Si l'intervention englobe le renforcement des capacités des États membres, par exemple l'aptitude à modifier la législation ou à préparer des plans de développement intégrant la norme, l'analyse documentaire devrait examiner l'expérience internationale dans ce domaine, y compris les études récentes sur la nature du développement des capacités, comme celle de Baser et Morgan (2008).

	gouvernements au niveau national à mettre en œuvre un instrument normatif.	
3. <i>Auto-évaluations</i>	Utiles au moment d'apprécier les connaissances et les compétences acquises par des personnes qui ont reçu une formation, un coaching ou un <i>mentoring</i> dans le cadre du travail normatif. Dans l'idéal, devraient être menées avant et après la formation, le coaching ou le <i>mentoring</i> .	Souvent assorties d'un biais positif, de sorte que les évaluateurs devraient se procurer une deuxième source de données factuelles, comme les superviseurs des participants à la formation.
4. <i>Étalonnage</i>	Potentiellement utile pour comparer les politiques générales, la législation ou les plans de développement de chaque gouvernement à un point de référence ou à un exemple idéal qui satisfait à tous les critères de l'ONU pour la norme ou le standard.	Nécessité de reconnaître les circonstances atténuantes dans certains pays, comme une faible capacité de rédaction d'instruments législatifs, ou le manque de volonté politique, qui pourraient rendre le point de référence difficile à atteindre.

76. Lorsque l'évaluation se concentre sur l'expérience de l'intégration du standard dans un ou quelques pays, il serait utile à l'évaluateur de préparer une chronologie recensant les principales interventions et les mesures prises par chaque pays pour accorder leurs politiques, lois et réglementations au standard. Cette chronologie²⁷ pourrait aider l'évaluateur à mettre en corrélation les données ayant trait aux changements qui ont eu lieu dans toutes les interventions normatives.

Tableau 11. Méthodes d'évaluation pour les initiatives visant à aider les gouvernements à mettre en œuvre des lois, des politiques et des plans de développement sur la base d'un standard international

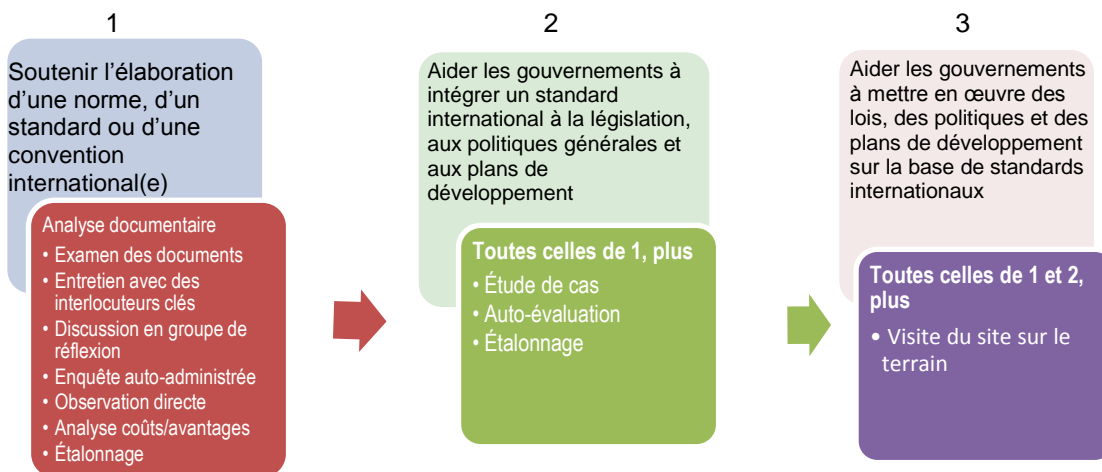
Méthodologie	Avantages	Inconvénients
1. <i>Toutes celles recensées dans les Tableaux 9 et 10 peuvent s'appliquer</i>		
2. <i>Visites des sites sur le terrain</i>	Les visites des sites sur le terrain donnent à l'évaluateur l'occasion de découvrir les points de vue des détenteurs de droits/bénéficiaires finaux. Elles peuvent aider l'évaluateur à valider les informations obtenues d'autres sources et grâce à d'autres méthodes, et elles peuvent aussi l'aider à comprendre directement les contraintes, comme le manque de fonds ou de compétences techniques,	Coûteuses lorsque de nombreux pays sont concernés.

²⁷ On peut trouver un exemple d'utilisation d'une chronologie et d'un étalonnage associé (voir article 4, Tableau 10) dans l'Évaluation de 2007 de la FAO des travaux de la Convention internationale pour la protection des végétaux et de ses arrangements institutionnels, disponible sur : <[ftp://ftp.fao.org/docrep/fao/meeting/012/k0233e02.pdf](http://ftp.fao.org/docrep/fao/meeting/012/k0233e02.pdf)>

	<p>auxquelles se heurtent les gouvernements ou les institutions nationaux lorsqu'ils tentent de mettre en œuvre des standards internationaux. La plupart des évaluations des Nations Unies englobent actuellement les visites sur le terrain dans leur méthodologie.</p>	
--	--	--

77. La Figure 3 fournit un résumé visuel rapide des méthodologies les plus courantes pour chacun des trois types d'évaluations de travaux normatifs dans le système des Nations Unies.

Figure 3. Méthodologies courantes par type d'évaluation



78. L'évaluateur n'est pas limité à ces méthodes courantes ; les méthodes spéciales décrites dans le Tableau 12 ci-dessous peuvent être utilisées de manière efficace pour évaluer des travaux normatifs dans les circonstances décrites.

Tableau 12. Méthodes d'évaluation à utiliser dans des circonstances spéciales

<p>Cartographie des effets²⁸</p>	<p>La cartographie des effets est un système de mesure conçu par le Centre de recherches pour le développement international (CRDI) au Canada. Au lieu de mener des appréciations des produits d'une intervention (p.ex. changements au niveau des politiques générales), il se concentre principalement sur les changements de comportement des personnes et organisations touchées par l'intervention. La cartographie des effets établit une vision des améliorations sur les plans humain, social et environnement à laquelle le projet espère contribuer, puis elle concentre le suivi et l'évaluation sur des facteurs et des acteurs à l'intérieur de sa sphère d'influence. La cartographie des effets accorde moins d'attention aux progrès réels de l'intervention qu'à l'influence de l'intervention (intentionnelle ou non) durant sa mise en œuvre. Elle convient surtout au suivi et à l'évaluation d'initiatives aidant les gouvernements à intégrer un standard international à la législation, aux politiques générales et aux plans de développement et pour aider les gouvernements à mettre en œuvre des lois, politiques et plans de développement sur la base d'un standard international.</p>
<p>Collecte des effets²⁹</p>	<p>La collecte des effets peut être appliquée dans le cadre d'évaluations de travaux normatifs lorsqu'il n'y a pas de effets prédéterminés sous forme d'un modèle logique ou d'une théorie du changement, ni d'informations de suivi des performances par rapport à des indicateurs prédéfinis. La méthode fonctionne dans le sens inverse des évaluations standard : l'évaluateur commence par « collecter » des effets – positifs et négatifs – à partir de rapports, d'entretiens et d'autres sources, puis procède à rebours pour déterminer la mesure dans laquelle le projet ou l'initiative ont contribué aux effets. Les effets collectés passent par un processus de sélection durant lequel des sources indépendantes et bien informées les valident et les justifient. Cette méthodologie convient dans des environnements normatifs extrêmement complexes dans lesquels la cause et l'effet ne sont pas évidents. Elle peut demander beaucoup de temps et requiert un degré important de participation. Surtout applicable pour évaluer les 2^e et 3^e types de travaux normatifs.</p>
<p>Reconstitution de processus</p>	<p>La reconstitution de processus est une méthode dont l'objectif est de générer et d'analyser des données qualitatives sur les mécanismes ou processus causaux, les événements, les actions, les attentes et autres variables qui interviennent et lient les causes hypothétiques aux effets observés³⁰. Oxfam UK a obtenu de bons résultats en utilisant la reconstitution de processus pour évaluer l'influence sur les politiques générales en Afrique du Sud et en Bolivie (Oxfam UK, 2012). Tout particulièrement adapté pour l'évaluation de travaux normatifs aidant les gouvernements à intégrer un standard international à la législation, aux politiques</p>

²⁸ Pour un complément d'informations sur la cartographie des effets, voir : www.idrc.ca/EN/Resources/Publications/Pages/ArticleDetails.aspx?PublicationID=1004

²⁹ L'Agence suisse pour le développement et la coopération a eu recours à la collecte des effets pour une évaluation du Programme mondial 2007-2011 BioNET, qui soutient un réseau environnemental mondial volontaire qui promeut une taxonomie standard liée à la mise en œuvre de la Convention sur la diversité biologique. Voir : www.outcomemapping.ca/resource/resource.php?id=374 pour un complément d'informations sur la collecte des effets.

³⁰ Voir : Collier, D. (2011). « Understanding Process Tracing ». *PS: Political Science and Politics*, 44:4. Berkeley : University of California, pp. 823-830. Tiré de polisci.berkeley.edu/people/faculty/CollierD/Understanding%20Process%20Tracing.pdf

	générales et aux plans de développement.
Analyse des systèmes	L'analyse des systèmes est l'étude d'ensembles d'entités en interaction et des relations entre elles. La méthode est relativement nouvelle dans le domaine de l'évaluation et elle peut se révéler utile pour analyser des systèmes complexes, comme ceux qui ont trait à la gouvernance, au commerce et à divers systèmes sociaux qui comportent des variables multiples, en reliant les relations et la communication qui peut influencer les résultats normatifs. Elle convient pour les trois catégories de travaux normatifs.
Changement le plus significatif	La méthode du changement le plus significatif (CPS) est une forme de suivi et d'évaluation participatifs qualitative et fondée sur les récits. Elle peut être utilisée pour évaluer les interventions liées aux droits de l'homme et à l'aide à la démocratie dans lesquelles les effets sont difficiles à prédire et dans lesquelles il peut ne pas y avoir d'indicateurs prédéterminés (ONUDI et PNUE, 2008). Le processus suppose la collecte de récits de changements significatifs émanant du terrain et la sélection systématique des plus significatifs de ces récits par des groupes de parties prenantes ou de membres du personnel désignés. La méthode CPS peut identifier les changements positifs et négatifs entraînés par une intervention. Elle peut aider les parties prenantes à concentrer leurs activités sur les moteurs du changement. Tout particulièrement adaptée à l'appréciation des effets et de l'impact des efforts fournis par les gouvernements nationaux pour appliquer les politiques, les lois et les plans de développement fondés sur des normes internationales au niveau communautaire.
Exploration positive (<i>Appreciative Inquiry</i>) ³¹	Méthodologie utilisée au départ dans le cadre du développement organisationnel et maintenant dans le domaine de l'évaluation, qui se base sur les forces d'une organisation et non sur ses défauts. L'exploration positive (AI) demande aux participants de s'interroger mutuellement en utilisant des questions qui génèrent des exemples, des histoires et des métaphores sur les aspects positifs de la vie de l'organisation. Les participants analysent les résultats en groupes en cherchant les thèmes et les sujets qui peuvent devenir la fondation du développement organisationnel positif. L'AI peut être appliquée dans le cadre d'évaluations participatives qui formulent les questions de l'évaluation sous un angle positif et qui attendent des participants qu'ils déterminent des actions pratiques pour produire un état positif souhaité. L'AI peut être utilisée pour évaluer les trois catégories de travaux normatifs.
Photovoice ³²	Photovoice est une méthode qui peut être appliquée dans le cadre des évaluations pour lesquelles les points de vue des jeunes, des groupes marginalisés ou privés de leur autonomie sont importants. Les participants sont priés de représenter leur point de vue en prenant des photographies, en parlant d'elles entre eux, en développant des récits pour accompagner leurs photos et en entreprenant un

³¹ Pour un complément d'informations sur l'AI, voir : [<appreciativeinquiry.case.edu/>](http://appreciativeinquiry.case.edu/)

³² Pour un complément d'informations, voir : [<en.wikipedia.org/wiki/Photovoice>](http://en.wikipedia.org/wiki/Photovoice)

	travail de proximité ou autre. Cette méthode pourrait être employée, par exemple, pour comprendre les points de vue des bénéficiaires de travaux normatifs qui ne répondront peut-être pas aux enquêtes et aux entretiens. Elle convient tout particulièrement à l'évaluation au niveau communautaire des effets et de l'impact des politiques générales, des lois et des plans de développement basés sur des normes internationales.
--	--

79. Il y a beaucoup d'autres méthodes, outils et ressources à la disposition des responsables de l'évaluation pour les évaluations de types particuliers de travaux normatifs. Par exemple, les responsables de l'évaluation pourraient utiliser l'Évaluation rurale participative³³ s'ils devaient évaluer des interventions normatives au niveau communautaire. Pour les évaluations de niveau national, des jeux de données longitudinaux mondiaux pourraient se révéler utiles, en particulier lorsque le travail normatif présente des problématiques de gouvernance. Freedom House, Polity et Transparency International figurent parmi les organisations qui disposent de données de ce type, et l'ONU elle-même a de nombreux indices potentiellement utiles, dont l'Indice de développement humain du PNUD. La mise en garde pour leur utilisation concerne la causalité : bien que les performances d'un pays pendant qu'il reçoit le soutien de l'ONU aient pu être améliorées selon l'indice, il est peu probable que cette amélioration ait été causée par le soutien de l'ONU à lui seul.

80. Un certain nombre de méthodes spéciales sont disponibles pour évaluer les initiatives de plaidoyer. Parmi elles on peut citer le suivi médiatique, la cartographie des politiques générales, la cartographie des réseaux et les entretiens structurés avec des « baromètres »³⁴. Ces méthodes, entre autres, sont décrites dans *User's Guide to Advocacy Evaluation Planning*, par Coffman (2009). Coffman soutient que l'évaluation du plaidoyer demande un état d'esprit non traditionnel, parce que le travail de plaidoyer efficace fluctue continuellement, s'ajustant aux changements qui surviennent dans l'environnement. Dans ce domaine de travail normatif, les effets planifiés sont susceptibles d'évoluer rapidement. Pour Coffman, « Le véritable défi consiste à évaluer ce qui survient en cours de route et les enseignements que l'on peut tirer de ce voyage » (Coffman, 2007, p. 32). Comme on le note dans le Chapitre 1, cela peut demander une approche non traditionnelle de l'évaluation, comme l'évaluation développementale.

81. *Evaluating Democracy Support* (Burnell, 2007) fait la chronique des nombreux défis à relever dans le cadre de l'évaluation des programmes qui soutiennent la démocratisation,

³³ L'Évaluation rurale participative est une méthodologie participative dotée d'un ensemble d'outils participatifs qui est souvent utilisée dans le domaine du développement rural pour veiller à ce que les connaissances et les opinions des parties prenantes locales figurent en bonne place dans la planification, la gestion et l'évaluation des programmes de développement.

³⁴ Les « baromètres » sont des jauges de tendances futures ou des prédicteurs d'événements futurs. Dans ce contexte, les baromètres sont des leaders d'opinion bien informés et innovants dont les points de vue sur les questions de politiques générales ont un poids et une valeur prédictive considérables. Dans le cadre de cette méthodologie, les baromètres sont informés avant l'entretien de ce que ce dernier couvrira de manière générale mais il ne leur est pas donné de détails précis pour veiller à ce que leurs réponses soient spontanées.

l'État de droit, la lutte contre la corruption et les droits de l'homme. Publié par l'Institut international pour la démocratie et l'assistance électorale (International IDEA) et l'Agence suédoise de coopération internationale au développement, ce document présente des arguments robustes en faveur des approches participatives et en faveur d'un mélange de mesures quantitatives et qualitatives.

82. L'évaluation de produits et de services normatifs comme les publications, les cours de formation, les bases de données, les réseaux de connaissances, et ainsi de suite, peut nécessiter des méthodes spéciales conjuguées à d'autres méthodes recensées dans les Tableaux 9, 10 et 11. Par exemple, les examens par des pairs peuvent se révéler utiles pour évaluer la qualité technique des supports normatifs. L'analyse Web est utile pour mesurer l'utilisation des ressources basées sur le Web au fil du temps. Lors de l'évaluation de produits normatifs, les responsables de l'évaluation souhaiteront peut-être les évaluer par rapport à des pratiques internationalement reconnues, comme la réalisation d'évaluations des besoins, les efforts pour faire intervenir les parties prenantes à leur conception, le ciblage des supports pour des publics différents, la mise à l'épreuve sur le terrain des supports avant de commencer à les utiliser, l'intégration des principes relatifs à l'égalité des sexes et aux droits de l'homme, et la garantie de budgets adéquats pour leur promotion et leur distribution.³⁵

Tâche 2.6 : Sélectionner l'équipe de l'évaluation

83. La sélection de l'équipe de l'évaluation fait partie des tâches les plus importantes que doit effectuer le responsable de l'évaluation. Bien que beaucoup d'aspects dépendent de l'axe de l'évaluation, de sa portée et de son budget, il convient de garder à l'esprit les points généraux suivants.

84. Tous les membres de l'équipe devraient bien connaître les normes, les standards ou les conventions au cœur de l'intervention. Des spécialistes dans certains sujets seront peut-être requis selon la nature du travail normatif. Si, par exemple, le travail est extrêmement technique, comme dans les domaines de la propriété intellectuelle, du droit international, de la biologie marine, de la sécurité aérienne ou de questions liées aux échanges commerciaux, un spécialiste peut se révéler extrêmement utile. Si le travail normatif englobe des activités de plaidoyer, il faudrait que l'équipe englobe une personne dotée de connaissances techniques spécialisées en matière d'évolution des politiques générales, de relations avec les médias, de communication dans le domaine du développement et, éventuellement, de médias sociaux. L'aspect le plus important est que les membres de l'équipe, y compris le chef d'équipe, doivent avoir des antécédents robustes en matière d'évaluation et une expérience large. L'aptitude à écrire de façon claire et concise est souvent une compétence négligée mais qui revêt une importance cruciale pour l'évaluation. Une évaluation peut s'effondrer si le chef d'équipe et ceux qui contribuent au rapport de l'évaluation ont des compétences d'écriture médiocres. Plus l'évaluation est complexe – et les évaluations des travaux

³⁵ Le modèle d'évaluation figurant dans l'Annexe 8 de l'évaluation des activités de développement des capacités de la FAO en Afrique (FAO, 2010) saisit nombre de ces critères d'évaluation.

normatifs sont généralement complexes – plus il est essentiel d’avoir de bonnes compétences d’écriture et d’analyse.

85. La plupart des évaluations de travaux normatifs de l’ONU examinées aux fins de ce manuel ont eu recours à des équipes hybrides comportant des membres internes et externes. Ceci dit, il peut parfois être avantageux d’utiliser seulement des évaluateurs internes. Dans d’autres cas, comme celui d’évaluations de travaux normatifs de haut niveau couvrant des agences tout entières ou bien lorsque des connaissances techniques hautement spécialisées sont requises, il peut se révéler nécessaire de faire appel à des experts externes.
86. Le responsable de l’évaluation doit être très sensible au genre et aux droits de l’homme, ainsi qu’aux nuances culturelles et politiques, au moment de composer l’équipe de l’évaluation. Dans l’idéal, les équipes engloberont un spécialiste de l’égalité des sexes et des droits de l’homme, et elles présenteront une parité femmes-hommes appropriée. Les consultants locaux peuvent compléter le travail des consultants internationaux chaque fois que des missions sur le terrain sont menées. Les connaissances sur la localité et les langues locales revêtent une importance cruciale parce que la plupart des évaluations de travaux normatifs dans le système de l’ONU nécessitent une compréhension des contextes social, organisationnel et institutionnel de l’intervention. Dans les situations qui demandent des modèles participatifs d’évaluation, le gouvernement et d’autres parties prenantes peuvent prendre part à la collecte et à l’analyse des données.
87. Des interprètes et des traducteurs sont indispensables dans les pays où il existe des barrières linguistiques considérables. Il convient de recruter les meilleurs interprètes et traducteurs que l’on pourra trouver, et de ne pas accepter les ressources proposées par les groupes de parties prenantes afin d’éviter tout filtrage indu des informations. Une fois recruté, l’interprète devient un membre à part entière de l’équipe. L’interprète a besoin de temps au début de l’évaluation pour découvrir le contexte normatif et pour se familiariser avec la terminologie technique de l’évaluation. Afin de faciliter le travail de l’interprète, le responsable de l’évaluation doit entraîner les autres membres de l’équipe dans l’art de mener des entretiens avec des phrases courtes.
88. Le responsable de l’évaluation devrait veiller à ce que le nombre de membres de l’équipe reste gérable pour tenter de contourner les problèmes inévitables qui affectent les grosses équipes et pour contrôler le coût de l’évaluation. Cela peut être accompli en sélectionnant comme membres de l’équipe des personnes dotées de compétences et d’expérience multiples et complémentaires. Le rôle de chaque membre de l’équipe doit être bien défini, tout comme les lignes de redevabilité. Le responsable de l’évaluation devrait prévoir du temps au début de l’évaluation pour veiller à ce que les membres de l’équipe soient mis au courant des aspects éthiques de l’évaluation.

Tâche 2.7 : Préparer le cahier des charges de l’évaluation

89. Le cahier des charges (CC) pour l’évaluation de travaux normatifs devrait suivre le même modèle que le CC pour l’évaluation d’initiatives de développement. Il devrait contenir, au minimum, la finalité, l’utilisation, la portée, l’axe central, les questions clés, les thèmes transversaux, la conception, la méthodologie, le calendrier de l’évaluation, ainsi que des

informations de référence sur l'intervention à évaluer, ainsi qu'une théorie du changement. Le CC devrait aussi décrire les rôles et les responsabilités des principaux participants à l'évaluation. Le document devrait décrire les normes, les standards ou les conventions de manière assez détaillée. Souvent, les CC des évaluations font référence aux lignes directrices éthiques et aux standards des Nations Unies à suivre. Certains CC stipulent quelles sont les connaissances techniques spécialisées requises au sein de l'équipe d'évaluation. Le CC devrait expliquer brièvement aux lecteurs pourquoi l'évaluation a lieu, sur quoi elle va se concentrer et qui y prendra part. Il n'englobe généralement pas la matrice de conception, qui est la base du plan pour mener l'évaluation. Dans l'idéal, le CC ne fera pas plus de 10 pages et il sera écrit dans un style simple pour pouvoir être largement distribué, auquel cas les informations budgétaires seront exclues.

Les membres de l'équipe prenant part à une évaluation de pays de la FAO, qui englobait une évaluation du travail normatif de la FAO, ont fait l'objet de nombreuses critiques légères concernant le CC de l'évaluation, qui s'étendait sur environ 30 pages. Certaines parties prenantes ont demandé que le rapport final soit plus court que le cahier des charges.

90. Le CC devrait rester une version préliminaire jusqu'à ce que l'évaluation commence. Cela fait partie des bonnes pratiques d'ajuster les questions et la méthodologie à la lumière de nouvelles informations reçues après que l'évaluation a commencé. Parfois, les consultants recrutés pour faire partie de l'équipe de l'évaluation ou les membres du groupe de référence apportent de nouveaux points de vue qui justifient l'apport de modifications au CC.

Encadré 7. Enseignements tirés de la préparation des évaluations de travaux normatifs

La participation de parties prenantes internes et, dans la mesure du possible, de parties prenantes externes pertinentes, tout au long de l'évaluation, accroît la pertinence perçue de cette évaluation, ainsi que son appropriation par les parties prenantes (UNEG, 2010).

« Il est important d'arriver à un juste milieu entre la promotion de l'appropriation des constatations et des recommandations de l'évaluation sans compromettre l'indépendance de l'évaluation, ce qui peut être facilité en définissant clairement le rôle d'un groupe de référence ou consultatif de ce type avant de commencer le travail » (UNEG, 2010, p. 4).

Il peut être difficile de préparer des indicateurs, en particulier pour certains types de travaux normatifs. D'après l'expérience, la meilleure manière d'effectuer cette tâche est au sein d'un groupe qui englobe des spécialistes du thème et une personne dotée de robustes antécédents de gestion axée sur les résultats et d'une expérience en préparation d'indicateurs.

Les rapports femmes/hommes sont un indicateur insuffisant pour l'égalité des sexes. Ce ne sont pas les chiffres, mais la substance de ces chiffres, qui compte – p. ex. représentation vs. participation (OIT, 2012b).

Concentrez l'évaluation sur ce qui est le plus important et sur ce qui est faisable. Dans les évaluations de travaux normatifs, plus on grimpe dans l'échelle des résultats, plus l'analyse deviendra difficile.

Ce qui compte ne peut pas toujours être compté et ce qui peut être compté ne compte pas forcément. (Albert Einstein)

Une réflexion rigoureuse l'emporte sur des méthodes rigoureuses. (M. Patton, 2011)

Encadré 8. Contrôle qualité – préparer les évaluations de travaux normatifs

- ✓ La finalité de l'évaluation du travail normatif a-t-elle été clairement formulée et l'utilisation des données de l'évaluation a-t-elle été déterminée ?
- ✓ Plusieurs questions clés ont-elles été identifiées comme axe de l'évaluation du travail normatif ?
- ✓ Les diverses options de conception ont-elles été considérées et une conception appropriée a-t-elle été sélectionnée ?
- ✓ Une analyse des parties prenantes a-t-elle été effectuée et une stratégie pour le travail de proximité a-t-elle été préparée ?
- ✓ A-t-on identifié des parties prenantes (femmes et hommes) pour participer à l'évaluation du travail normatif et leurs rôles et responsabilités ont-ils été décrits ?
- ✓ Des sources clés de données ont-elles été identifiées ?
- ✓ Plusieurs méthodes ont-elles été considérées et une combinaison des méthodes les plus appropriées a-t-elle été sélectionnée à la lumière de la finalité et du contexte de l'évaluation ?
- ✓ Des indicateurs appropriés et sensibles au genre ont-ils été sélectionnés pour chaque résultat et impact ?
- ✓ La conception et les méthodes sélectionnées prennent-elles en compte les questions clés de l'évaluation ?
- ✓ La conception et les méthodes permettent-elles une analyse exhaustive des questions transversales, y compris les droits de l'homme et l'égalité des sexes ?
- ✓ Des solutions ont-elles été examinées pour les défis prévus ?

Ressources utiles

Burnell, P. (éd.) (2007). *Evaluating Democracy Support. Methods and Experiences*. Stockholm : Institut international pour la démocratie et l'assistance électorale (International IDEA) et Agence suédoise de coopération internationale au développement.

Agence canadienne de développement international (1997). *Guide to Gender-Sensitive Indicators*. Disponible sur : <www.acdi-cida.gc.ca/>

Coffman, J. (2009). *User's Guide to Advocacy Evaluation Planning*. Cambridge, Massachusetts: Harvard Family Research Project.

Coffman, J. (2007). "What's Different About Evaluating Advocacy and Policy Change?" *The Evaluation Exchange Vol. XIII, No. 1*.

Pawson, R. et Tilley, N. (1997). *Realistic Evaluation*. Thousand Oaks, Californie : Sage.

Reisman, J., Gienapp, A., et Stachowiak, S. (2007). *A Guide to Measuring Advocacy and Policy*. Baltimore : Annie E. Casey Foundation.

Stern, E., Stame, N., Mayne, J., Forss, K, Davies R. et Befani, B. (2012). “Broadening the Range of Designs and Methods for Impact Evaluations.” *Department for International Development Working Paper No. 38*. Londres : Department for International Development.

ONUSIDA (2011). *An Introduction to Indicators*. Disponible sur :

www.unaids.org/en/media/unaids/contentassets/documents/document/2010/8_2-Intro-to-IndicatorsFMEF.pdf

UNEG (2005). *Norms and Standards for Evaluation in the UN System*. Disponible sur :

www.uneval.org

Wilson-Grau, R. et Britt, H. (2012). *Outcome Harvesting*. Le Caire : Ford Foundation, Middle East and North Africa Office. Disponible sur :

www.outcomemapping.ca/resource/resource.php?id=374

Banque mondiale (2010). *Cost-benefit Analysis in World Bank Projects: Overview*. New York: Groupe d'évaluation indépendante, Banque mondiale. Disponible sur :

siteresources.worldbank.org/INTOED/Resources/cba_overview.pdf

Initiative internationale pour l'évaluation de l'impact (International Initiative for Impact Evaluation) : Importante collection de ressources de référence, de trousseaux à outils, de documents téléchargeables et d'autres ressources sur la méthodologie et les meilleures pratiques pour l'évaluation de l'impact : www.3ieimpact.org/

NONIE : Un réseau de plus de 100 bureaux d'évaluation internationaux, y compris l'ONU, la Banque mondiale, d'autres organisations de développement, des pays en développement et diverses organisations régionales et mondiales.

Communication Initiative Network : Informations complètes sur la conception, la mise en œuvre et l'évaluation de campagnes de communication et de plaidoyer : www.comminet.com

Chapitre 3. Collecter et analyser les données

Tâche 3.1 : Collecter les données

91. Ce manuel se tourne maintenant vers certains des défis courants associés à l'évaluation du travail normatif, laquelle supposera dans la plupart des cas une appréciation beaucoup plus qualitative que quantitative.

« La plupart des activités de plaidoyer qui ont lieu lors de réunions, de réceptions et de dîners privés, loin des projecteurs, ne sont pas documentées. Nombre des activités de plaidoyer de haut niveau ne le sont pas non plus. Certains événements sont brièvement décrits dans des rapports et dans les médias, mais le détail des tactiques employées dans le cadre du plaidoyer ne l'est pas. Par exemple, il y a relativement peu de documentation qui décrit les années d'exercice de pression intensif qui ont abouti à l'Ordonnance relative à la justice pour mineurs... L'absence de données de ce type entrave l'apprentissage organisationnel. »

- *Évaluation de la Stratégie de plaidoyer et de mobilisation sociale du gouvernement du Pakistan/de l'UNICEF* (Stiles, 2002, p. 17)

92. Souvent le plus grand défi consiste à se procurer des données fiables pour évaluer les effets (sur le plan des comportements et des changements institutionnels) dérivés du travail normatif. L'équipe de l'évaluation devra rassembler les informations portant sur la situation avant et après l'intervention à partir d'un examen des documents, d'entretiens et d'autres méthodes. Parmi les sources évidentes d'informations figurent les agents chargés de l'intervention, les partenaires clés chargés de la mise en œuvre et les autres parties prenantes clés. Il ne faut pas oublier les représentants de l'ONU qui ont entrepris des démarches formelles et informelles. Certains des meilleurs aperçus – souvent les pépites d'or de la perception – peuvent émaner de sources surprenantes : membres du personnel de terrain subalternes, fonctionnaires gouvernementaux à la retraite, chefs d'entreprise, journalistes et même chauffeurs.

93. Il est essentiel de recueillir les opinions des femmes ainsi que des hommes, parce que leurs besoins sont distincts et l'impact des interventions normatives sur chaque sexe est souvent différent. Plusieurs évaluations de l'ONU (par exemple, OIT 2012a) examinées pour ce manuel ont constaté que le renforcement des rôles stéréotypés pour les femmes et les hommes dans les travaux normatifs réduit les opportunités économiques pour les femmes. Les évaluations ont constaté que les travaux normatifs qui n'ont pas englobé une analyse rigoureuse des questions de genre peuvent engendrer des effets négatifs pour les femmes et les enfants, même lorsque les femmes et les enfants sont ciblés comme bénéficiaires clés (UNICEF, 2008).

94. L'analyse de l'influence des questions externes sur les effets du travail normatif est essentielle, mais c'est aussi l'un des aspects les plus difficiles de l'évaluation. Certaines parties prenantes peuvent rechigner à reconnaître que le travail de l'ONU a influencé leurs décisions d'adopter des normes internationales en dépit de preuves du contraire. L'évaluateur doit tenir compte des différentes manières de voir la réalité en triangulant les données émanant de sources multiples et de champs d'enquête multiples au moment de juger du mérite d'une intervention. De temps en temps, l'évaluateur trouvera peut-être un signe révélateur d'influence, comme une référence à une norme de l'ONU dans un énoncé de politique générale, ou l'utilisation répandue d'un outil analytique de l'ONU. Les

« Les questions externes devraient être minutieusement analysées afin de comprendre comment elles interagissent avec un projet précis. C'est là l'élément le plus difficile de l'évaluation : les influences externes peuvent être nombreuses et difficiles à prévoir, en particulier dans les programmes de nature politique visant la démocratie. Par ailleurs, le manque de consensus théorique entre les experts, les universitaires et les praticiens sur la relation causale entre les variables et leurs effets politiques plus larges vient compliquer l'analyse. »

- Burnell (éd.) (2007), p. 107

évaluateurs doivent aussi aller au-delà des effets et indicateurs prescrits, car les résultats inattendus pourraient être plus importants que les ceux qui étaient prévus.

95. Les gens sont au cœur même des évaluations et, souvent, c'est traiter avec eux qui constitue le plus grand défi à relever dans le cadre d'évaluations normatives de nature qualitative. Souvent, lors de l'évaluation d'un travail normatif, les évaluateurs doivent gérer les attentes, les rapports de pouvoir, les opinions politiques personnelles et les idées arrêtées sur le rôle de l'évaluation. Un examen des évaluations qui se sont mal passées suggère que la bonne communication est un ingrédient essentiel du succès (Gervais, 2010). La communication autour des standards internationaux est délicate en elle-

même parce que les gens ont des interprétations variées des standards et il peut parfois leur déplaire de devoir s'y plier. Expliquer la finalité de l'évaluation et la manière dont les informations seront utilisées relève des bonnes pratiques.

96. Que devrait faire le responsable de l'évaluation si l'équipe de l'évaluation ne reçoit pas de réponses aux questions de l'évaluation ? Bien que cela arrive rarement s'il y a eu une bonne planification, le responsable de l'évaluation devrait avoir un plan d'urgence. La meilleure stratégie consiste parfois à faire le point après la première série d'entretiens ou après la première étude de cas de pays, et à déterminer s'il est possible de tirer des enseignements et si l'équipe de l'évaluation peut trouver un moyen de recentrer l'évaluation. Si l'équipe n'y parvient pas, le mieux est de s'adresser à nouveau au groupe de référence pour lui demander conseil, et d'apporter les changements nécessaires à la conception et aux méthodes de l'évaluation³⁶.

Tâche 3.2 : Analyser les données

97. La nature du travail normatif des Nations Unies soulève des défis pour l'analyse des données, comme l'illustre le rapport sur une évaluation de la programmation d'UNIFEM pour faciliter la mise en œuvre de la CEDAW en Asie du Sud-Est :

Bien que notre analyse ait mis en évidence des signes de progrès par rapport aux produits associés et par rapport aux indicateurs liés aux effets pris séparément, il reste encore à déterminer si et comment ces réussites démontrent à elles toutes une volonté politique accrue et un engagement généré/renforcé en faveur de la mise en œuvre de la CEDAW. (UNIFEM, 2008, p. 9)

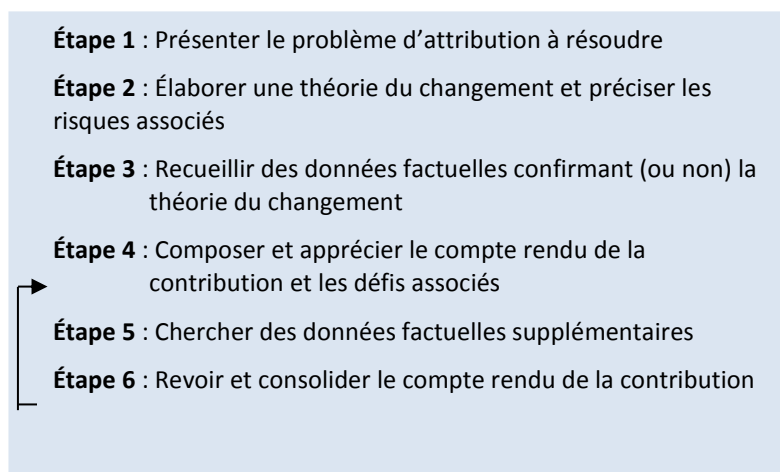
³⁶ Les écrits récents sur l'évaluation pourraient peut-être orienter sur la marche à suivre lorsqu'une évaluation se passe mal. Voir, par exemple, le numéro spécial du *Canadian Journal of Program Evaluation*, Vol. 25, No. 3, 2010.

98. Que donnent les éléments recueillis à eux tous ? Telle est l'énigme pour l'évaluateur. Cette question est tout particulièrement importante dans le cadre de l'évaluation de travaux normatifs lorsqu'il incombe aux gouvernements, en tant que garants d'obligations, de mettre en œuvre des normes, des standards et des conventions internationaux, tandis que des organisations de l'ONU jouent un rôle de facilitation et/ou de soutien. Pour résoudre cette question, l'évaluateur doit faire la distinction entre *attribution* et *contribution*, concepts définis comme suit :

L'attribution fait intervenir une thèse causale sur l'intervention comme cause de l'impact, et une mesure de la part de l'impact qui peut être associée à l'intervention. Cela se distingue de la « contribution », qui avance une thèse causale sur la question de savoir si et comment une intervention a contribué à un impact observé. Il s'agit en général de vérifier une théorie du changement en tenant compte d'autres facteurs influents, ce qui réduit l'incertitude quant à la contribution apportée par l'intervention (Stern et al., 2012).

99. À mesure que l'on remonte la chaîne logique vers l'impact, il est parfois plus exact de parler de contribution ou d'influence plutôt que de lien direct de cause à effet (UNEG, 2012). Lorsque des travaux normatifs de l'ONU ont influencé les politiques générales ou la législation, « ce que l'on peut attendre de mieux, c'est une interprétation plausible, et non une « preuve » irréfutable » (Stern et al., 2012, p. 34).
100. Le travail mené par Mayne (2011) sur l'analyse de la contribution fournit aux évaluateurs du travail normatif de l'ONU une approche pratique pour avancer des thèses causales et réduire le caractère incertain de la causalité lorsque les données sont limitées et lorsqu'il n'est pas possible d'effectuer des conceptions expérimentales ou quasi-expérimentales. L'analyse de la contribution est un processus en six étapes (voir Encadré 9) qui consiste à vérifier et à revoir au besoin la théorie du changement de l'intervention, en accordant une attention particulière aux facteurs qui influent sur les effets et l'impact.

Encadré 9. Étapes d'une analyse de la contribution



Source : Adapté de Mayne, 2011.

101. L'analyse de la contribution est fondée sur la prémisse selon laquelle une intervention fait partie d'un bloc causal plus large. Lorsque l'intervention donne de bons résultats, elle constitue un élément essentiel du bloc, et lorsque le bloc donne de bons résultats, cela suffit à entraîner le changement prévu. Il s'ensuit que l'intervention à elle seule ne suffit pas à entraîner le changement, mais qu'elle est un élément essentiel du bloc causal, auquel cas elle est une cause *contributive*. L'intervention est considérée particulièrement robuste lorsqu'elle déclenche le changement. Lorsqu'il y a suffisamment de preuves pour arriver à cette conclusion, l'évaluateur peut affirmer que ce bloc causal complet a suffi à entraîner les résultats collectifs et que les actions de l'intervention ont été le facteur déclenchant qui ont donné lieu à ce résultat (Mayne, 2011).
102. L'analyse de la contribution oblige l'évaluateur à chercher les facteurs directs et indirects qui influencent les changements dans des environnements de programmation complexes courants dans le travail normatif de l'ONU et à expliciter les suppositions sous-jacentes. Non seulement l'analyse de la contribution est utile, elle est aussi suffisamment rigoureuse pour inspirer confiance dans les conclusions des évaluateurs en ce qui concerne l'efficacité des interventions normatives.
103. Les évaluateurs doivent se montrer vigilants au moment d'employer un prisme basé sur les droits de l'homme et l'égalité des sexes pour leur analyse des données. Cela vaut pour l'évaluation de tous les travaux normatifs, pas seulement les travaux expressément axés sur les droits de l'homme ou l'égalité des sexes. L'analyse doit être appliquée à trois niveaux : la conception originale de l'intervention pour déterminer si les préoccupations relatives aux droits de l'homme et à l'égalité des sexes ont été prises en compte ; la mesure dans laquelle l'intervention a ensuite véritablement appliqué les principes relatifs aux droits de l'homme et à l'égalité des sexes et/ou une approche basée sur les droits de l'homme ; et une analyse des effets et de l'impact d'une intervention pour déterminer la mesure dans laquelle ces effets et cet impact sont alignés sur les principes relatifs aux droits de l'homme et à l'égalité des sexes et la façon dont ils pourraient toucher

différemment les femmes et les hommes. Le concours de spécialistes des droits de l'homme et du genre peut se révéler des plus précieux au moment de déterminer ces divers aspects. Pour des informations plus détaillées, le guide de l'UNEG *Intégrer les droits de l'homme et l'égalité des sexes aux évaluations* est utile. Des ressources supplémentaires sont recensées à la fin de ce chapitre.

104. Mise en garde concernant les cibles : certains rapports d'évaluation examinés dans le cadre de la préparation de ce manuel fondaient leurs évaluations en grande mesure sur la mesure dans laquelle les cibles ont été satisfaites. Aucun n'a pris la peine de déterminer si les cibles étaient réalistes ou si d'autres priorités peut-être plus importantes s'étaient dégagées durant le travail normatif. Au mieux, les cibles racontent une partie de l'« histoire des performances » et, surtout dans le cadre du travail normatif, les cibles sont susceptibles de changer à tout moment. Bien que les cibles doivent être fréquemment ajustées, les responsables et les bailleurs de fonds rechignent souvent à les modifier une fois qu'elles ont été fixées. C'est pour ces raisons que l'évaluateur devrait les examiner minutieusement avant de tirer des conclusions hâtives.
105. L'utilisation de matrices pour catégoriser les informations peut faciliter l'analyse des données³⁷. Un atelier de réflexion peut également être un moyen participatif d'analyser les données. Certaines évaluations de travaux normatifs ont organisé un atelier immédiatement après avoir recueilli des données, dans le cadre duquel les membres de l'équipe d'évaluation reviennent sur le cahier des charges et examinent de près chaque question de l'évaluation. L'atelier doit être bien facilité pour garantir une participation et une réflexion critique complètes. Il est bon d'avoir recours à des fiches pour faciliter l'identification et l'organisation des points clés de manière visuelle. Cette méthode est tout particulièrement utile au moment de trier des volumes énormes de données pour identifier ce qui importe véritablement. Elle constitue aussi un bon moyen de préparer la présentation des constatations préliminaires de l'équipe au groupe de référence ou aux parties prenantes à la fin d'une mission d'évaluation.

Tâche 3.3 : Identifier les enseignements et les pratiques les mieux adaptées

106. Durant leur analyse de leurs données, les membres de l'équipe d'évaluation devraient prendre le temps d'identifier les enseignements et les pratiques les mieux adaptées. Les termes « enseignements » et « pratiques les mieux adaptées » sont utilisés délibérément. Les enseignements ne deviennent pas des « leçons apprises » avant d'être véritablement mis en œuvre. Certaines organisations reproduisent les mêmes erreurs alors qu'elles affirment avoir appris leurs leçons. L'expression « pratiques les mieux adaptées » convient mieux que « meilleures pratiques », parce que toutes les pratiques, quelle que soit leur

³⁷ Une évaluation indépendante du Programme pour une production plus propre de l'ONUDI et du PNUE en 2008 a eu recours à des matrices pour noter les performances par rapport aux conseils en matière de politiques générales dans 18 pays. Les évaluateurs ont organisé les données sur les produits, les effets et les impacts dans trois catégories : échelle des résultats, type de données factuelles et robustesse des données (ONUDI et PNUE, 2008, pp. 92-93).

efficacité dans un contexte particulier, ne conviennent pas à tous les contextes. Ces nuances sont tout particulièrement applicables à l'évaluation du travail normatif du fait des contextes variés dans lesquels le travail a lieu. Il s'ensuit que les enseignements figurant à la fin de chacun des chapitres de ce manuel peuvent ne pas être applicables à chaque évaluation de travail normatif.

107. Il est souvent difficile d'identifier les enseignements et les pratiques les mieux adaptées : ce qui figure dans certains rapports d'évaluation est soit insuffisamment ancré dans des données factuelles, soit écrit sous forme de recommandations. Les enseignements émanent souvent d'enquêtes qualitatives qui posent la question : « Pourquoi ceci est-il arrivé et quels sont les enseignements que l'on peut en tirer ? ». Les pratiques les mieux adaptées découlent aussi largement d'enquêtes qualitatives qui posent la question : « Pourquoi ceci a-t-il bien fonctionné et peut-on l'appliquer à des travaux normatifs similaires dans des conditions analogues ? ». La tâche consistant à identifier les enseignements et les pratiques les mieux adaptées devrait être répartie au sein de l'équipe d'évaluation. Il est important de prévoir au moins une journée complète pour cette tâche, parce qu'elle requiert souvent une réflexion et une discussion considérables. La contribution de parties prenantes clés peut être utile. On pourrait demander aux personnes les plus étroitement associées à l'intervention dès le stade de la délimitation de l'évaluation d'identifier elles-mêmes des enseignements et certaines des pratiques les mieux adaptées, que l'équipe d'évaluation pourrait ensuite examiner de plus près durant le stade de la collecte de données (au lieu de tomber dessus par hasard vers le milieu du processus). Au minimum, il faudrait donner aux parties prenantes l'occasion d'examiner et de faire des commentaires sur les enseignements et les pratiques les mieux adaptées une fois qu'elles ont été rédigées.
108. Comme la plupart des enseignements sont assimilés par essais et erreurs, les évaluateurs doivent chercher à obtenir ces informations dans le cadre d'entretiens car les responsables sont susceptibles d'hésiter à parler de travaux qui ne se passent pas bien. Certains enseignements et pratiques les mieux adaptées proviendront éventuellement d'analyses documentaires. Par exemple, on trouve des enseignements et des exemples de pratiques les mieux adaptées extrêmement utiles dans les documents ayant trait à la formation du consensus, au plaidoyer, à la communication en vue de changements de comportement, au travail en réseaux, au développement des capacités et à l'exercice d'influence sur les politiques générales.

Encadré 10. Enseignements en matière de collecte et d'analyse des données

Il est tout particulièrement important d'accorder une attention aux effets imprévus potentiels pour les évaluations de travaux normatifs, comme dans le cas où une convention est en vigueur depuis quelque temps, mais sans guère de suivi, compte tenu du fait que, dans les environnements complexes, tout ou presque peut arriver sans être directement contrôlé par l'agence qui mène les travaux normatifs.

Il faudrait également identifier les voies causales multiples et parfois simultanées. Par exemple, le plaidoyer peut supposer une collaboration directe avec le gouvernement, et fréquemment aux échelons politiques élevés, en même temps que le développement d'un soutien au sein de l'administration. Souvent, il fait intervenir la mobilisation de la communauté. Il peut englober une action directe et/ou indirecte, tant formelle – une campagne médiatique – qu'informelle, y compris des discussions avec des fonctionnaires de haut niveau.

La coopération des partenaires est nécessaire pour la collecte des données.

« L'obtention d'informations auprès de femmes ainsi que d'hommes peut accroître le coût et le temps requis pour la collecte des données. Cet aspect doit être considéré, mais est souvent négligé durant la planification et la budgétisation de l'exercice de S&E [suivi et évaluation]. » (OIT, 2012b, p. 9)

« La meilleure façon de répondre aux questions d'EI [évaluation de l'impact] sur les effets d'une intervention suppose d'adopter un prisme « contributif » : les programmes couronnés de succès font partie d'un « bloc causal » suffisant – ils ne fonctionnent pas isolément. » (Stern *et al.*, 2012, p. 36)

Encadré 11. Contrôle qualité – collecte et analyse des données

- ✓ Les données de référence ont-elles été identifiées ou recréées ?
- ✓ Les informations contextuelles ont-elles été examinées de près ?
- ✓ Les données ont-elles été consignées soigneusement et de manière qui réduisent au minimum le biais éventuel ?
- ✓ Les membres de l'équipe d'évaluation ont-ils pris part à l'examen et l'analyse des données collectées ?
- ✓ Les évaluateurs ont-ils eu recours à la triangulation dans leur analyse des données ?
- ✓ Une analyse rigoureuse de la contribution a-t-elle été effectuée qui identifie le bloc causal, détermine si le travail normatif était suffisant et nécessaire pour entraîner le changement identifié et si le travail normatif a déclenché le changement ?
- ✓ Les constatations clés ont-elles été présentées aux parties prenantes clés et validées par elles ?

Ressources utiles

- Equitas – Centre international d'éducation aux droits humains et le Bureau du Haut-Commissariat des Nations Unies aux droits de l'homme (2011). *Évaluer les activités de formation aux droits de l'homme – Manuel destiné aux éducateurs dans le domaine des droits de l'homme*. Série sur la formation professionnelle No. 18. Disponible sur : <http://www.ohchr.org/Documents/Publications/EvaluationHandbookPT18.pdf>
- Burnell, P. (éd.) (2007). *Evaluating Democracy Support. Methods and Experiences*. Stockholm: Institut international pour la démocratie et l'assistance électorale (International IDEA) et Agence suédoise de coopération internationale au développement. Disponible sur : http://www.idea.int/publications/evaluating_democracy_support/index.cfm
- Conseil International sur les politiques des droits humains (2012). « No Perfect Measure: Rethinking Evaluation and Assessment of Human Rights Work. A Report of a Workshop. » Genève : ICHRP.
- Institut international pour la démocratie et l'assistance électorale (2002). *Handbook on Democracy Assessment*. La Haye : Kluwer Law. Disponible sur : http://www.idea.int/publications/sod/democracy_assessment.cfm
- Organisation internationale du travail (2012). *Integrating Gender Equality in Monitoring and Evaluation of Projects. Guidance Note 4*. Genève : Organisation internationale du travail, Unité d'évaluation.
- CAD/OCDE. 2010. *Normes de qualité pour l'évaluation du développement*. Disponible sur : <http://www.oecd.org/dataoecd/55/0/44798177.pdf>
- Patton, M. (2001). *Qualitative Research & Evaluation Methods, 3rd edition*. Thousand Oaks, Californie : Sage.
- Mayne, J. (ed) (2012). « Contribution Analysis. » *Evaluation*, Numéro spécial, 18(3).
- Mayne, J. (2011). « Contribution Analysis: Addressing Cause and Effect » dans *Evaluating the Complex*, K. Forss, M. Marra et R. Schwartz (éd.). Piscataway, New Jersey : Transaction Publishers.
- White, H. et Phillips, D. (2012). *Addressing attribution of cause and effect in small impact evaluations: towards an integrated framework*. New Delhi: International Initiative for Impact Evaluation.
- Portail sur l'égalité des sexes d'ONU Femmes : <http://genderevaluation.unwomen.org/>

Chapitre 4. Préparer le rapport d'évaluation

Tâche 4.1 : Préparer le rapport d'évaluation

109. Un rapport final bien écrit est le « joyau de la couronne » d'une évaluation réussie. Les constatations et les recommandations de l'évaluation ont plus de chances d'être prises au sérieux et d'être mises en application si le rapport est bien structuré et agréable à lire. Il existe de nombreux guides utiles pour faciliter la rédaction de rapports, dont certains sont indiqués à la fin de ce chapitre. L'aide-mémoire relatif à la qualité figurant à la fin du chapitre est un rappel utile des points clés se dégageant de nombre de ces guides.
110. Étant donné qu'aucun manuel de rédaction de rapports n'a été rédigé exclusivement pour l'évaluation du travail normatif, les suggestions suivantes se fondent sur un examen de rapports d'évaluation dans le cadre de la préparation de ce manuel.
111. Il est essentiel, lorsque l'on décrit des travaux normatifs, de fournir un contexte, comme la toile de fond sociopolitique et culturelle et l'historique du travail. Il est également important de citer les contributions d'autres organisations, tant locales qu'internationales. Les lecteurs doivent comprendre l'environnement dans lequel le travail normatif a eu lieu, y compris ses opportunités, ses défis, et les éléments majeurs qui en influencent les effets. Ils doivent par ailleurs connaître les moteurs sous-jacents à tous changements normatifs. Ces informations feront la lumière sur l'analyse de la contribution et faciliteront la formulation des recommandations. On trouvera d'excellents exemples de couverture contextuelle dans l'évaluation DevInfo de l'UNICEF (UNICEF, 2009) et l'Étude de performance globale de 2010 du FEM.
112. Le responsable de l'évaluation devrait veiller à ce que le rapport saisisse non seulement les effets prévus mais aussi les effets imprévus, positifs ainsi que négatifs. Étant donné la nature non linéaire de certains travaux normatifs, comme la recherche d'un accord sur un standard, le plaidoyer, le dialogue autour de politiques générales, le développement des capacités et le renforcement de la confiance, quelques changements imprévus sont inévitables, et certains pourraient être plus significatifs que ceux qui avaient été planifiés.
113. Les conclusions et les recommandations d'une évaluation revêtent une importance cruciale. Prenez le temps de les rédiger avec soin, car, si elles sont mal écrites, elles peuvent se transformer en point d'ignition de critiques qui pourraient éclipser tous les éléments positifs du reste du rapport. N'oubliez pas que les recommandations doivent être destinées aux personnes dotées de la responsabilité de les mettre en œuvre. Dans la plupart des évaluations de travaux normatifs de l'ONU, ces personnes seraient les responsables de l'ONU et leurs homologues et partenaires au sein des gouvernements. En cas de doute, testez les recommandations auprès de membres du groupe de référence et d'autres parties prenantes avant que le rapport ne soit finalisé. Au moment de formuler des recommandations, suggérez des manières de renforcer les initiatives normatives, certes, mais n'oubliez pas de recommander, lorsque c'est justifié, des mesures visant à améliorer le suivi et l'évaluation de travaux normatifs futurs. La plupart des évaluations de l'OIT

examinées aux fins de ce manuel, par exemple, ont bien préparé et orienté les recommandations.

114. Les évaluations ne nécessitent pas toutes des recommandations. Par exemple, l'évaluation de l'impact du travail de l'ONUDI sur les standards, la méthodologie, les essais et la qualité au Sri Lanka (ONUDI, 2010) ne comportait aucune recommandation, mais se concentrait plutôt sur une série d'enseignements visant à orienter les travaux futurs.
115. Il est aussi important de reconnaître les limites de l'évaluation, en particulier lorsque les données de référence ne sont pas disponibles et que le suivi de l'information est limité. Les contraintes auxquelles on se heurte au moment de mener l'évaluation devraient figurer dans la section du rapport portant sur la méthodologie, ainsi qu'une description de ce qui a été fait pour composer avec ces limites.

Encadré 12. Enseignements – Préparer le rapport d'évaluation

Un rapport écrit en langage simple et sans jargon est beaucoup plus susceptible d'être lu.

La meilleure manière de veiller à ce qu'un rapport soit lu est de supprimer les parties que peu de personnes voudront lire.

« Je parviens souvent au bon degré de spontanéité dans mon écriture à la cinquième mouture. » — John Kenneth Galbraith

« La brièveté est l'âme de l'esprit. » — William Shakespeare, *Hamlet*

Encadré 13. Contrôle qualité – Préparer le rapport d'évaluation

- ✓ Y a-t-il suffisamment de données factuelles (provenant de sources et de champs d'enquête multiples) pour soutenir chaque constatation principale ?
- ✓ Les conclusions sont-elles bien justifiées par les données factuelles présentées et logiquement connectées aux constatations de l'évaluation ?
- ✓ Les limites de l'évaluation et sa méthodologie sont-elles décrites ?
- ✓ Les jugements évaluatifs sont-ils raisonnables et justifiés au vu des données factuelles présentées et du contexte normatif ?
- ✓ Les jugements exprimés dans le rapport sont-ils équilibrés et justes ?
- ✓ Les recommandations sont-elles justifiées, ciblées et possibles à appliquer ?
- ✓ Les recommandations traduisent-elles une compréhension de l'organisation et des contraintes potentielles pour le suivi ?
- ✓ Le rapport explique-t-il comment l'égalité des sexes et une approche basée sur les droits de l'homme ont été intégrées dans l'évaluation, et ces thèmes transversaux se reflètent-ils dans les constatations, les conclusions et les recommandations ?
- ✓ Le rapport utilise-t-il constamment un langage sensible au genre et basé sur les droits de l'homme, y compris des données ventilées par sexe, âge et handicap ?
- ✓ Le rapport respecte-t-il la vie privée ?
- ✓ Y a-t-il un résumé d'orientation qui exprime de manière succincte la raison d'être de l'évaluation, sa méthodologie et ses principales constatations, conclusions et recommandations ?

Ressources utiles

Art and Architecture of Writing Evaluation Reports: Secrétariat du Conseil du Trésor du Canada (2004). Disponible sur : www.tbs-sct.gc.ca/cee/career-carriere/workshops-ateliers/aawer-amrre-eng.pdf

UNEG Quality Checklist for Evaluation Reports (2010). Disponible sur : www.unevaluation.org/papersandpubs/documentdetail.jsp?doc_id=607

UNEG Règles d'évaluation applicables au sein du système des Nations Unies (2005). Disponible sur : www.unevaluation.org/unevaluationstandards

Chapitre 5. Donner suite à l'évaluation

Tâche 5.1 : Distribuer le rapport d'évaluation

116. Bien que chaque organisation de l'ONU ait sa propre politique générale concernant la distribution des rapports d'évaluation, le responsable de l'évaluation peut prendre des mesures pour en améliorer la distribution et l'utilisation. Ces dernières doivent faire l'objet d'une réflexion lors de la planification de l'évaluation, et non être mises de côté jusqu'au dernier moment. Dans l'idéal, elles devraient être abordées dans la stratégie de travail de proximité de l'évaluation, comme on le mentionne au Chapitre 2. Une planification précoce est tout particulièrement importante pour les évaluations de travaux normatifs dont l'un des objectifs premiers est l'apprentissage. Il ne faut pas confondre distribution et utilisation ; la distribution peut faciliter l'utilisation, mais ce n'est pas la même chose que prendre des décisions, améliorer la programmation ou générer des connaissances (Patton, 2008). Il est également important de garder à l'esprit les différents publics auxquels le rapport d'évaluation est destiné. Bien que les responsables de l'ONU soient souvent le public premier, les évaluations de travaux normatifs visant à aider les gouvernements à mettre en œuvre la législation, les politiques générales et les plans de développement sur la base de normes et standards internationaux peuvent comporter des recommandations destinées aux représentants gouvernementaux, auquel cas la stratégie de distribution devrait englober une session de suivi avec le gouvernement et éventuellement des activités de suivi continues.
117. Comme on l'a mentionné dans les chapitres antérieurs, une fois que le rapport parvient aux responsables supérieurs, dans l'idéal ils sauront d'ores et déjà ce qu'il contient, puisqu'ils ont joué un rôle dans le groupe de référence de l'évaluation et/ou dans d'autres modalités de débriefing. Il relève d'une bonne stratégie d'éviter tout élément de surprise dans les rapports d'évaluation, en particulier si leur contenu risque d'être controversé.
118. Les standards de l'UNEG recommandent aux organisations d'avoir une politique explicite de divulgation qui assure « la diffusion transparente des résultats de l'évaluation, notamment en rendant les rapports largement accessibles aux organes directeurs et au public, sauf dans les cas où il faut assurer à certaines des parties prenantes un degré raisonnable de protection et de confidentialité » (UNEG, 2005a, p. 4).
119. En signe de respect et conformément aux pratiques recommandées par l'UNEG, un exemplaire du rapport, une fois approuvé, sera remis à ceux qui ont participé à l'évaluation. Dans un esprit de transparence, ceux qui ont pris le temps, malgré leurs emplois du temps chargés, de fournir des informations, devraient avoir l'occasion de voir les résultats. L'inclusion du rapport final sur le site Web de l'organisation fait désormais partie des pratiques courantes. Des exemplaires du rapport doivent aussi être mis à la disposition de tous à travers les bureaux régionaux et nationaux des Nations Unies. Lorsqu'il existe des barrières linguistiques, le responsable de l'évaluation fera traduire le résumé d'orientation afin qu'il puisse être distribué de manière large.

120. Dans les situations où le rapport d'évaluation comporte des recommandations pour améliorer le travail normatif en cours, et lorsque les parties prenantes locales sont réceptives à des améliorations, le responsable de l'évaluation devrait envisager une mission de suivi pour discuter des recommandations et de leur mise en œuvre. Bien que cela ne fasse pas partie des pratiques standard, c'est un aspect qui a le potentiel de renforcer l'appropriation, l'apprentissage et l'adoption des recommandations des évaluateurs au niveau local.

Tâche 5.2 : Préparer une réponse de la part des équipes de direction

121. L'UNEG encourage les organisations membres à se doter de « mécanismes appropriés de suivi de l'évaluation » pour veiller à ce que les « recommandations issues des évaluations soient utilisées comme il se doit et appliquées en temps utile, les constatations de l'évaluation devant servir aux activités futures » (UNEG, 2005a, p. 4). Les normes de l'UNEG attendent de l'organe directeur et de l'équipe de direction qu'ils répondent aux recommandations de l'évaluation et des organisations membres qu'elles effectuent « un suivi systématique de la mise en œuvre des recommandations de l'évaluation que l'équipe de direction et/ou l'organe directeur ont acceptées... », ainsi qu'un « ... compte rendu périodique de la situation de la mise en œuvre... » (UNEG, 2005b, p. 10).

Tâche 5.3 : Utiliser l'évaluation

122. Les mécanismes formels de la réponse et du suivi de l'équipe de direction ont pour objectif de veiller à ce que les recommandations soient appliquées, mais ils ne constituent pas une garantie. Des mesures supplémentaires sont souvent requises pour tirer le meilleur parti des connaissances acquises lors de l'évaluation et des enseignements générés dans le cadre de ce processus.

123. Étant donné la probabilité d'une demande accrue d'évaluations de travaux normatifs au sein du système des Nations Unies, les responsables de l'évaluation devraient envisager d'échanger leurs rapports d'évaluation entre eux et de mettre sur pied une communauté de pratique au sein de la famille de l'ONU. Les informations peuvent dans un premier temps être échangées de façon informelle mais, si l'intérêt s'accroît, cela pourrait se transformer en réseau de connaissances soutenu par un référentiel électronique. Une partie des enseignements pourraient être incorporés dans des programmes de développement professionnel destinés aux responsables de l'évaluation à l'avenir. Le partage d'une base de données de consultants compétents en évaluation des travaux normatifs pourrait être un autre dérivé constructif.

124. Les évaluations de travaux normatifs revêtiraient un intérêt pour les cercles professionnels de l'évaluation, qui sont maintenant nombreux. Les bureaux de l'ONU chargés de l'évaluation pourraient nouer des liens avec eux afin de renforcer les pratiques professionnelles.

Tâche 5.4 : Évaluer l'évaluation

125. Cette tâche, rarement effectuée, a une valeur potentielle considérable. Le responsable de l'évaluation pourrait tirer des enseignements en demandant un *feedback* (retour) sur l'évaluation quelques mois, ou plus longtemps, après. Les constatations ont-elles été utilisées ? Quels sont les aspects de l'évaluation que les parties prenantes ont trouvés utiles ? Qu'est-ce qui a bien fonctionné et moins bien fonctionné quant à la conception, la méthodologie et le processus de mise en œuvre de l'évaluation ? Quels enseignements peut-on tirer de l'expérience pour améliorer l'évaluation des travaux normatifs à l'avenir ? On pourrait répondre rapidement à certaines de ces questions à partir des informations archivées ; d'autres pourraient être examinées au moyen d'une téléconférence, d'une courte enquête, d'échanges d'e-mails ou de réunions. Ces activités de suivi ne sont pas forcément coûteuses. Lorsque c'est justifié, les responsables de l'évaluation pourraient entreprendre des examens par des pairs d'évaluations de travaux normatifs afin de favoriser l'apprentissage et de renforcer la manière dont ces évaluations sont entreprises dans l'ensemble du système des Nations Unies.

Encadré 14. Enseignements – suivi de l'évaluation

Préparez la voie à l'utilisation dès la planification de l'évaluation.

Identifiez un champion du changement de niveau supérieur qui puisse donner suite aux enseignements et aux recommandations à partir de l'évaluation du travail normatif.

Ressources utiles

Costa, P. (2012). *Study on Communicating Evaluation Results*. Paris : Réseau DevCom de l'OCDE. Disponible sur : www.oecd.org/dac/evaluation/StudyonCommunicatingEvaluationResults_FINAL_091112.pdf

UNEG (2005). *Normes d'évaluation applicables dans le système des Nations Unies*. Disponible sur : www.uneval.org/normsandstandards/index.jsp

UNEG (2005). *Règles d'évaluation applicables au sein du système des Nations Unies*. Disponible sur : www.uneval.org/normsandstandards/index.jsp

Documents examinés

- Bamberger, M. et Rugh, J. (2008). « A Framework for Assessing the Quality, Conclusions, Validity and Utility of Evaluations. » Document présenté lors de la Conférence de l'American Evaluation Association, Baltimore, 2007.
- Baser, H. et Morgan, P. (2008). *Capacity, Change and Performance*. Maastricht : Centre européen de gestion des politiques de développement.
- Bolger, J. et Stiles, M. (2011). *End-of-project Evaluation of the Southeast Asia Regional Cooperation in Human Development (SEARCH) Project*. Gatineau : Southeast Asia Regional Programme, Agence canadienne de développement international.
- Burnell, P. (ed.) (2007). *Evaluating Democracy Support. Methods and Experiences*. Stockholm: Institut international pour la démocratie et l'assistance électorale (International IDEA) et Agence suédoise de coopération internationale au développement. Disponible sur : www.idea.int/publications/evaluating_democracy_support/index.cfm
- Coffman, J. (2009). *User's Guide to Advocacy Evaluation Planning*. Cambridge, Massachusetts : Harvard Family Research Project.
- Coffman, J. (2007). « What's Different About Evaluating Advocacy and Policy Change? » *The Evaluation Exchange* Vol. XIII, No. 1.
- Collier, D. (2011). « Understanding Process Tracing » *PS: Political Science and Politics*, 44:4. Berkeley : University of California, pp. 823 -830. Disponible sur : polisci.berkeley.edu/people/faculty/CollierD/Understanding%20Process%20Tracing.pdf
- Costa, P. (2012). *Study on Communicating Evaluation Results*. Paris : Réseau DevCom de l'OCDE.
- Coulibaly, K. (2010). *Better Work Baselineing and Impact Measurement Tools and Methodologies*. Genève : Organisation internationale du travail.
- Dozois, E., Langois, M. et Blanchet-Cohen, N. (2010). *DE.201.A Practitioner's Guide to Developmental Evaluation*. Montréal : The J. W. McConnell Family Foundation.
- El-Zanati et associés (2009). *Djibouti Communication for Development Survey (C4DS)*. Djibouti : Fonds des Nations Unies pour l'enfance.
- Equitas – Centre international d'éducation aux droits humains et le Bureau du Haut-Commissariat des Nations Unies aux droits de l'homme. (2011). *Évaluer les activités de formation aux droits de l'homme – Manuel destiné aux éducateurs dans le domaine des droits de l'homme*. Série sur la formation professionnelle No. 18.
- Fetterman, D. (2001). *Foundations of Empowerment Evaluation*. Thousand Oaks, CA: Sage.

- Organisation des Nations Unies pour l'alimentation et l'agriculture (2010). *Évaluation des activités de la FAO en matière de renforcement des capacités en Afrique*. PC104/5. Rome : Bureau de l'évaluation.
- Organisation des Nations Unies pour l'alimentation et l'agriculture (2009). *Independent Evaluation of FAO Corporate Objective B-1: "International instruments concerning food, agriculture, fisheries and forestry, and the production, safe use and fair exchange of agricultural, fishery and forestry goods."* Rome : Bureau de l'évaluation, FAO.
- Organisation des Nations Unies pour l'alimentation et l'agriculture (2008). *Évaluation indépendante du rôle et des activités de la FAO dans le domaine statistique*. Rome : Organisation des Nations Unies pour l'alimentation et l'agriculture.
- Organisation des Nations Unies pour l'alimentation et l'agriculture (n. d.). « Normative and Operational Activities. » Présentation PowerPoint. Rome : FAO.
- Funnell, S. et Rogers, P. (2011). *Purposeful Program Theory*. San Francisco : Jossey-Bass.
- Gardner, A. et Geierstanger, S. (2007). « Working With Logic Models to Evaluate Policy and Advocacy Program. » *The Evaluation Exchange XIII (1)*.
- Gervais, M. (2010). « As I Recall—Or How to Take Advantage of Less-Than-Successful Evaluation Experiences. » *The Canadian Journal of Program Evaluation*, 25(3).
- Gill, A. et Gilbert, N., (2011). *Written Evidence to the House of Commons Committee on Science and Technology and references therein*.
- Fonds pour l'environnement mondial (2010). *OPS4: Progrès vers la réalisation d'un impact. Quatrième Étude de la performance globale du FEM. Version exécutive*. Washington, DC : Bureau d'évaluation du Fonds pour l'environnement mondial.
- GrantCraft (2006). *Mapping Change: Using Theory of Change to Guide Planning and Evaluation*.
- Guidance Note on Framing and Conducting Impact Evaluations of Normative Work by the UN System.
- Conseil International sur les politiques des droits humains (2012). "No Perfect Measure: Rethinking Evaluation and Assessment of Human Rights Work. A Report of a Workshop." Genève : ICHRP.
- Institut international pour la démocratie et l'assistance électorale (International IDEA) (2002). *Handbook on Democracy Assessment*. La Haye : Kluwer Law.
- Organisation internationale du travail (2012a). *Time-Bound Programme - Urban Informal Economic Programme on Prevention and Elimination of the Worst Forms of Child Labour in the Urban Informal Economy of Dhaka Metropolitan Area. Final evaluation*. Genève:

Bureau international du travail, Programme international pour l'abolition du travail des enfants.

Organisation internationale du travail (2012b). *Integrating Gender Equality in Monitoring and Evaluation of Projects. Guidance Note 4*. Genève: Organisation internationale du travail, Unité d'évaluation.

Organisation internationale du travail (2012c). *Promoting freedom of association and collective bargaining rights in the rural and export processing sectors: August 2009-December 2011*. Rapport de l'évaluation finale indépendante. Genève: Bureau international du travail.

Organisation internationale du travail (2012d). *Support to the Indonesian Time-Bound Programme of the Elimination of Worst Forms of Child Labour-Phase II. Final Evaluation*. Genève: Bureau international du travail, Programme international pour l'abolition du travail des enfants.

Organisation internationale du travail (2012e). *Independent Final Evaluation Report of Project: Combating Forced Labour and Trafficking of Indonesian Migrant Workers, Phase II*. Genève : Unité d'évaluation.

Organisation internationale du travail (2011). *Consolidating the Legal and Institutional Foundations of Social Dialogue in the Countries of Western Balkans and Moldova*. Genève : Unité d'évaluation.

Organisation internationale du travail (2008). *Independent Evaluation of the ILO's Strategy to Support Member States to Improve the Impact of International Labour Standards*. Genève: Bureau international du travail.

Klugman, B. (2010). *Evaluating Social Justice Advocacy. A Values-Based Approach*. Washington DC : Centre for Evaluation Innovation.

Leeuw, F. et Vaessen, J. (2009). *Impact Evaluations and Development. Nonie Guidance on Impact Evaluation*. Washington, DC : La Banque mondiale.

Levine, C. et Chapoy, C. (2013). « Building on What Works: Commitment to Evaluation (c2e) Indicator. » *Working Paper 19*. New Delhi: International Initiative for Impact Evaluation.

Louie, J. et Guthrie, K. (2007). « Strategies for Assessing Policy Change Efforts: A Prospective Approach. » *The Evaluation Exchange XIII (1)*.

Mackie, J. (1974). *The Cement of the Universe*. Oxford: Oxford University Press.

Mayne, J. (ed.) (2012). « Contribution Analysis », *Evaluation*, Numéro spécial, 18(3).

Mayne, J. (2011) « Contribution Analysis: Addressing Cause and Effect » dans *Evaluating the Complex*, K. Forss, M. Marra et R. Schwartz (Eds.). Piscataway, New Jersey : Transaction Publishers.

- Morra-Imas, L. et Rist, R. (2009). *The Road to Results. Designing and Conducting Effective Development Evaluations*. Washington DC : la Banque mondiale.
- Muhweezi, A., Msham, H. et Mourshed, S. (2010). *Nile Transboundary Environment Action Project. Terminal Evaluation Report*. Entebbe : Nile Basin Initiative Secretariat.
- Nexus Associates, Inc. (2012). *Better Work Haiti. Mid-term Evaluation*. Genève : Organisation internationale du travail.
- Comité d'aide au développement/Organisation de coopération et de développement économiques (2012). *Efficacité du développement et de l'aide. Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats*. Paris : CAD/OCDE.
- Comité d'aide au développement/Organisation de coopération et de développement économiques (2010). *Normes de qualité pour l'évaluation du développement*. Disponible sur : www.oecd.org/dataoecd/55/0/44798177.pdf
- Comité d'aide au développement/Organisation de coopération et de développement économiques (2002). *Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats*. Paris: OCDE, Groupe de travail du CAD sur l'évaluation de l'aide.
- Ortiz, E. et Inomata, T. (2009). *Assessment of the Global Mechanism of the United Nations Convention to Combat Desertification (JIU/REP/2009/4)*. Genève: United Nations, Joint Inspection Unit.
- Oxfam UK (2012). *Process Tracing. Draft Protocol*. Guidance on Evaluating Policy Influencing Work. Disponible sur : www.oxfam.org.uk
- Patton, M. (2011a). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York : Guilford Press.
- Patton, M. (2011b). « Improving Rigor in Service-Learning Research. » Dans Hatcher, J. et Bringle, R. *Understanding Service-Learning and Community Engagement: Crossing Boundaries Through Research*. Nouvelle-Orléans : International Association for Service-Learning and Community Engagement.
- Patton, M. (2008). *Utilization-focused Evaluation: The New Century Text 4th edition*. Thousand Oaks, CA : Sage Publications.
- Pawson, R. et Tilley, N. (1997). *Realistic Evaluation*. Thousand Oaks et Londres : Sage.
- Quinn, M. et Clark, K. (2006). *Progressing the Implementation of CEDAW in the Pacific. An Independent Evaluation of UNIFEM's Programme*. New York : UNIFEM.
- Reisman, J., Gienapp, A., et Stachowiak, S. (2007). *A Guide to Measuring Advocacy and Policy*. Baltimore : Annie E. Casey Foundation.

- Rog, D. (1985). « A Methodological Assessment of Evaluability Assessments. » Thèse de doctorat, Vanderbilt University, Nashville, Tennessee.
- Roman-Morey, E. et Zahran, M. (2011). *Evaluation of the Scope, Organization, Effectiveness and Approach of the Work of the United Nations in Mine Action (JIU/REP 2011/11)*. Genève : Nations Unies, Unité d'inspection conjointe.
- Scriven, M. (1993). *Hard-won lessons in program evaluation. New Directions for Program Evaluation, Series No. 58*. San Francisco : Jossey Bass.
- Stern, E., Stame, N., Mayne, J., Forss, K, Davies R. et Befani, B. (2012). « Broadening the Range of Designs and Methods for Impact Evaluations. » Department for International Development Working Paper No. 38. Londres : Department for International Development (Département britannique pour le développement international).
- Stiles, M. (2011). *End of Project Evaluation of the Migrant Labour Rights Project (China)*. Gatineau, Québec : Agence canadienne de développement international.
- Stiles, M. (2009). *Evaluation of CIDA-funded Human Rights Cooperation between the University of Ottawa and Peking University*. Gatineau, Québec : Agence canadienne de développement international.
- Stiles, M. (2008). *An Assessment of CIDA's Contribution to Legal and Judicial Reform in China*. Gatineau : China Programme, Agence canadienne de développement international.
- Stiles, M. (2007). *An Evaluation of ESCAP's Approach to Capacity Development*. Bangkok : Commission économique et sociale des Nations Unies pour l'Asie et le Pacifique.
- Stiles, M. (2002). *Evaluation of the Government of Pakistan-UNICEF Advocacy and Social Mobilization Strategy, January 1999-June 2001*. Kathmandu : Fonds des Nations Unies pour l'enfance, Bureau régional d'Asie du Sud.
- Stuart, J. (2007). « Necessity Leads to Innovative Evaluation Approach and Practice. » *The Evaluation Exchange XIII (1)*.
- Tickner, V. (2006). *Impact Assessment of the FAO/MIC Marketing Management Assistance Project in Mozambique, financed by the EC, over the period 2001 – 2005*. Rome : Food and Agricultural Organization of the United Nations.
- UNIFEM (2008). UNIFEM Programme. *Facilitating CEDAW Implementation in Southeast Asia. Evaluation Report*. New York : UNIFEM Unité d'évaluation.
- UNIFEM (2004). *Assessment of From Global to Local: A Convention Monitoring and Implementation Project*. New York : UNIFEM, Innovation and Learning Unit.
- Fonds des Nations Unies pour l'enfance (2010a). *Advocacy Toolkit: A guide to influencing decisions that improve children's lives*. New York: UNICEF.

- Fonds des Nations Unies pour l'enfance (2010b). Full Terms of Reference: Consultancy Services for the Global Evaluation of the Application of a Human Rights Based Approach in UNICEF Programming.
- Fonds des Nations Unies pour l'enfance (2010c). *In Practice Newsletter: Refocusing on Equity*. UNICEF.
- Fonds des Nations Unies pour l'enfance (2010d). *In Practice Newsletter: Communication for Development*. UNICEF.
- Fonds des Nations Unies pour l'enfance (2010e). *Qualitative Study on Essential Family Practices for the prevention and integrated management of child illness and nutrition*. Bamako, Mali : UNICEF.
- Fonds des Nations Unies pour l'enfance (2009). *Global Evaluation of DevInfo*. New York: Bureau d'évaluation de l'UNICEF.
- Fonds des Nations Unies pour l'enfance (2008). *Evaluation of Gender Policy Implementation in UNICEF*. New York : Bureau d'évaluation de l'UNICEF.
- Fonds des Nations Unies pour l'enfance (n.d.). *Reintegration Programme for Children Associated With Armed Forces and Armed Groups Including Children Affected by Conflict in Nepal*. Note d'information.
- Commission économique et sociale des Nations Unies pour l'Asie de l'Ouest (2011). *Annual Report of Evaluations 2010*. Beyrouth : ESCWA.
- Conseil économique et social des Nations Unies (2011). *Programme Evaluation of the Department of Economic and Social Affairs. Report of the Office of Internal Oversight Services*, E/AC.51/2011/2. New York : Conseil économique et social des Nations Unies.
- Conseil économique et social des Nations Unies (2007). *Report of the Office of Internal Oversight Services on the in-depth evaluation of political affairs: Security Council affairs*, E/AC.51/2007/2/Add.2. New York : Conseil économique et social des Nations Unies.
- Organisation des Nations Unies pour l'éducation, la science et la culture (2013). L'évaluation du travail normatif de l'UNESCO dans le domaine de la culture (première partie) – Convention de 2003 pour la sauvegarde du patrimoine culturel immatériel. Paris: UNESCO.
- Organisation des Nations Unies pour l'éducation, la science et la culture (2012). *Évaluation de la phase pilote du Fonds international pour la diversité culturelle*. Paris : UNESCO.
- Organisation des Nations Unies pour l'éducation, la science et la culture (2012b). *Review of Findings on UNESCO's Normative Work in Culture Captured by Previous Evaluations*. Paris: UNESCO.

Organisation des Nations Unies pour l'éducation, la science et la culture (2010). *Évaluation de de l'Objectif stratégique de programme 11 de l'UNESCO, « Protéger et promouvoir de façon durable le patrimoine culturel »*. Paris: UNESCO.

Groupe des Nations Unies pour l'évaluation (2013). *Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management*. New York : UNEG.

Groupe des Nations Unies pour l'évaluation (2012). Terms of Reference. Evaluation of Normative Work in the UN System. Document non publié.

Groupe des Nations Unies pour l'évaluation, (2012). UNEG Task Force on Evaluation of Normative Work - Survey Analysis. New York : UNEG.

Groupe des Nations Unies pour l'évaluation (2011). *Integrating Human Rights and Gender Equality in Evaluation – Towards UNEG Guidance*. New York : UNEG.

Groupe des Nations Unies pour l'évaluation (2010a). *UNEG Good Practice Guidelines for Follow-up to Evaluations*. UNEG/G(2010)3. New York : UNEG.

Groupe des Nations Unies pour l'évaluation (2010b). *UNEG Quality Checklist for Evaluation Reports*, UNEG/G(2010)/2. New York: UNEG.

Groupe des Nations Unies pour l'évaluation (2005a). *Standards for Evaluation in the UN System*. New York : UNEG. Disponible sur : www.unevaluation.org/unevaluationstandards

Groupe des Nations Unies pour l'évaluation (2005b). *Norms for Evaluation in the UN System*. New York: UNEG. Disponible sur : www.uneval.org

Organisation des Nations Unies pour le développement industriel (2010a). *Evaluation UNIDO SMTQ [Standards, Metrology, Testing and Quality] Projects in Sri Lanka*. Vienne : Groupe d'évaluation de l'ONUDI.

Organisation des Nations Unies pour le développement industriel (2010b). *Thematic Evaluation Report. UNIDO activities in the areas of Standards, Metrology, Testing and Quality (SMTQ)*. Vienne : Groupe d'évaluation de l'ONUDI.

Organisation des Nations Unies pour le développement industriel et Programme des Nations Unies pour l'environnement (2008). *Évaluation indépendante ONUDI-PNUE du Programme de production plus propre*. Vienna : ONUDI.

Universal Management Group (2012). *Global Evaluation of the Application of a Human Rights Based Approach to UNICEF Programming*. New York : UNICEF.

Commission de réglementation nucléaire des États-Unis, (2005). *Assessment of the quality of the NRC Research Projects: Memorandum*.

Banque mondiale (2010). *Cost Benefit Analysis in World Bank Projects*. Washington, DC : Banque mondiale.

Banque mondiale (2004). *Suivi et évaluation : quelques outils, méthodes et approches*. Washington, DC : Département de l'évaluation des opérations, Banque mondiale.

White, H. et Phillips, D. (2012). *Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework*. New Delhi: International Initiative for Impact Evaluation.

Wilson-Grau, R. et Britt, H. (2012). *Outcome Harvesting*. Le Caire : Ford Foundation, Middle East and North Africa Office.