

# Critères d'évaluation

Greet Peersman

## CENTRE DE RECHERCHE DE L'UNICEF

Le Centre de recherche Innocenti est le bureau de l'UNICEF spécialisé en recherche scientifique. Les objectifs prioritaires du Centre de recherche sont d'améliorer la compréhension internationale des questions liées aux droits des enfants afin de faciliter la pleine application de la Convention relative à ces droits dans le monde entier. Le Centre a pour objectif de mettre en place un cadre intégré pour la gestion de la recherche et des connaissances au sein de l'organisation, afin de soutenir ses programmes et politiques à l'échelle mondiale. Les publications du Centre, en présentant un vaste éventail d'opinions, contribuent au débat international sur les questions liées aux droits des enfants.

Les opinions exprimées n'engagent que leurs auteurs et/ou réviseurs et sont publiées afin d'encourager le dialogue sur les méthodes d'évaluation d'impact. Cette publication ne reflète pas nécessairement les politiques ou points de vue de l'UNICEF.

## NOTES MÉTHODOLOGIQUES DU CENTRE DE RECHERCHE

Les notes méthodologiques du Centre de recherche de l'UNICEF ont le but de partager des pratiques, des méthodes et des modèles ainsi que des recommandations de chercheurs et analystes reconnus. Elles sont dirigées principalement au personnel de l'UNICEF qui conduit des recherches ou qui interprète les résultats et analyses des évaluations réalisées par des chercheurs externes pour les prises de décision en phase de programmation ou à l'appui des politiques ou de la sensibilisation.

Cette note méthodologique a suivi un procès interne de révision par pairs.

Ce texte n'a pas été édité conformément aux standards officiels de publication et UNICEF décline toute responsabilité pour les possibles erreurs.

La reproduction de parties de cette publication est permise à condition que la source soit dûment citée. Pour l'utilisation d'une partie substantielle ou de la totalité de la publication, veuillez bien vous adresser au Département de Communication au suivant courriel [florence@unicef.org](mailto:florence@unicef.org)

Pour consulter ou télécharger ces notes méthodologiques, visitez <http://www.unicef-irc.org/KM/IE/>

Nous conseillons d'utiliser la citation suivante pour toute référence au document présent :

Peersman, G. (2014). Critères d'évaluation, *Note méthodologique n° 3*, Centre de recherche Innocenti, Florence.

**Remerciements** : Cette note méthodologique a profité des conseils de beaucoup d'individus. L'auteur et le Centre de recherche désirent remercier tous ceux qui ont contribué et en particulier :

**Pour leur contribution** : Dugan Frasers Patricia Rogers

**Pour leur révision** : Nikola Balvin, Samuel Bickel, David Parker

© Fonds des Nations Unies pour l'enfance (UNICEF), septembre 2014

Centre de recherche Innocenti de l'UNICEF

Piazza SS. Annunziata, 12

50122 Florencia (Italia)

Tél. : (+39) 055 20 330

Fax : (+39) 055 2033 220

[florence@unicef.org](mailto:florence@unicef.org)

[www.unicef-irc.org](http://www.unicef-irc.org)

## 1. CRITÈRES D'ÉVALUATION : UNE BRÈVE DESCRIPTION

L'évaluation repose sur une combinaison de faits et de valeurs (principes, attributs ou qualités considérés comme intrinsèquement positifs, souhaitables, importants et ayant un intérêt général<sup>1</sup> comme celui « d'être juste pour tous ») visant à apprécier l'intérêt d'une intervention (un programme ou une politique). Les critères d'évaluation précisent les valeurs qui seront utilisées dans une évaluation. La présente note aborde spécifiquement leur utilisation dans les évaluations d'impact (études qui fournissent des informations sur les effets à long terme d'une intervention ; voir la Note n° 1, Présentation de l'évaluation d'impact), bien que ces critères puissent être utilisés dans différents types d'évaluation.

L'UNICEF utilise une palette de critères d'évaluation pour guider ses évaluations ; tous ne sont pas utilisés dans chacune des évaluations, n'étant parfois adaptés qu'à certains types d'évaluations ou d'interventions. Les [termes de référence](#) (TdR) de l'évaluation doivent préciser les critères qu'il conviendra d'utiliser.

Les évaluations d'impact des interventions de l'UNICEF<sup>2</sup> utilisent, entre autres, les critères d'évaluation suivants :

### Les critères généraux de l'OCDE-CAD<sup>3</sup>

- **Pertinence** : mesure selon laquelle les objectifs d'une intervention correspondent aux attentes des bénéficiaires, aux besoins du pays, aux priorités internationales et aux politiques des partenaires.
- **Efficacité** : mesure selon laquelle les objectifs de l'intervention ont été atteints, ou sont en train de l'être, compte tenu de leur importance relative.
- **Efficience** : mesure selon laquelle les ressources/intrants (fonds, expertise, temps, équipements, etc.) sont convertis en résultats de façon économe.
- **Impact** : effets à long terme, positifs et négatifs, primaires et secondaires, induits par une intervention, directement ou non, intentionnellement ou non.
- **Durabilité** : continuation des bénéfices résultant de l'intervention après l'interruption de l'aide au développement. Les interventions doivent être écologiquement et financièrement viables. Lorsque l'accent n'est pas mis sur l'aide extérieure, la durabilité peut être définie comme la capacité des principales parties prenantes à faire durer les bénéfices résultant de l'intervention – après l'arrêt du financement par les donateurs – en s'efforçant d'utiliser des ressources disponibles à l'échelle locale.

### Les critères relatifs à l'équité, à l'égalité entre les sexes et à l'adoption d'une approche de la programmation fondée sur les droits de l'homme

- **Équité** : principe fondamental de justice dans les processus décisionnels et leurs résultats. Pour l'UNICEF, cela implique que tous les enfants aient le droit de survivre, de se développer et de s'épanouir pleinement sans faire l'objet de discriminations, de préjugés ou de favoritisme.
- **Égalité entre les sexes** : pour l'UNICEF, cela signifie de promouvoir l'égalité des droits des femmes et des filles, et de soutenir leur pleine participation au développement politique, social et économique des communautés dans lesquelles elles vivent<sup>4</sup>.

<sup>1</sup> Stufflebeam, D. L., 2001, « Evaluation values and criteria checklist », Western Michigan University Checklist Project, Kalamazoo. Disponible à l'adresse : [https://www.wmich.edu/sites/default/files/attachments/u350/2014/values\\_criteria.pdf](https://www.wmich.edu/sites/default/files/attachments/u350/2014/values_criteria.pdf)

<sup>2</sup> Lorsque l'utilisation des critères d'évaluation dans le cadre de l'évaluation d'un programme ou d'une politique présente d'importantes variations, celles-ci seront spécifiquement mentionnées dans la présente note.

<sup>3</sup> Organisation de coopération et de développement économiques, Comité d'aide au développement, « Evaluation of development programmes, DAC Criteria for Evaluating Development Assistance », page Internet, OCDE, <http://www.oecd.org/fr/cad/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>

- **Approche de la programmation fondée sur les droits de l'homme** : les cinq principes fondamentaux sous-tendant cette approche peuvent être utilisés comme critères d'évaluation : la normativité, la non-discrimination, la participation, la transparence et la responsabilité<sup>5</sup>.

### Les critères d'évaluation<sup>6</sup> de l'aide humanitaire (utilisés en complément de certains ou de l'ensemble des critères susmentionnés)

- **Couverture** : la nécessité « d'atteindre des catégories importantes de la population victimes de souffrances qui mettent leur vie en danger, quel que soit le lieu où elles se trouvent, pour leur apporter une aide et une protection qui soient à la mesure de leurs besoins et exemptes de toute considération d'ordre politique inutile »<sup>7</sup>.
- **Coordination** : la nécessité pour les différentes parties prenantes d'une intervention d'urgence de coordonner leur action ; l'intervention d'une seule organisation ne peut être évaluée indépendamment, notamment parce que les mesures qui peuvent paraître appropriées du point de vue d'un acteur ne le seront pas forcément du point de vue du système dans son ensemble.
- **Protection** : la fourniture de services de sécurité et de protection aux membres de la population cible (p. ex. protection contre des éléments armés opérant dans la zone d'intervention, ou fourniture de services de sécurité à l'intérieur d'un camp de réfugiés ou de personnes déplacées) est essentielle pour garantir l'efficacité de l'action humanitaire et des secours.
- **Cohérence** (des politiques) : la nécessité de s'assurer que les politiques concernant l'aide humanitaire et les politiques en matière de sécurité, de développement, de commerce et de défense militaire sont compatibles.

Ces critères d'évaluation doivent être considérés comme des « concepts » devant être pris en compte lors de l'évaluation. Ils ne sont toutefois pas suffisamment précis pour être appliqués de façon systématique et transparente et permettre d'aboutir à des jugements évaluatifs sur les programmes ou les politiques. Pour chaque critère « générique », des critères plus spécifiques tels que des critères de référence et/ou des normes<sup>8</sup> – adaptés au type d'intervention et à son contexte – doivent être définis et convenus avec les principales parties prenantes.

À titre d'exemple, lors d'une évaluation mondiale de l'application par l'UNICEF de l'approche de la programmation fondée sur les droits de l'homme, les évaluateurs ont relevé différentes façons de concevoir cette approche. Ils ont mené des consultations et recherché un consensus afin de hiérarchiser cinq principes fondamentaux (tels qu'énumérés ci-dessus), puis ont défini des indicateurs clés pour évaluer les programmes spécifiques de l'UNICEF faisant l'objet d'une évaluation.

---

<sup>4</sup> Fonds des Nations Unies pour l'enfance, 2010, Préparer un avenir d'égalité, Politique de l'UNICEF sur l'égalité des sexes et l'autonomisation des filles et des femmes, UNICEF, New York

<sup>5</sup> Fonds des Nations Unies pour l'enfance, 2012, Évaluation mondiale de l'application par l'UNICEF de l'approche de la programmation fondée sur les droits de l'homme, Rapport final – Volume I, UNICEF, New York (à noter : un sixième principe d'indivisibilité/interdépendance a été considéré comme inhérent au principe de la normativité)

<sup>6</sup> Tirés du chapitre « Linking evaluation criteria to evaluation questions: Additional criteria for humanitarian assistance programmes » du *UNICEF Monitoring and Evaluation Training Resource*. Ce document est en cours de révision, c'est pourquoi aucun numéro de page n'a été spécifié et seul un extrait a été fourni.

<sup>7</sup> Organisation de coopération et de développement économiques, Comité d'aide au développement, 1999, *Conseils pour l'évaluation de l'aide humanitaire apportée dans les situations d'urgence complexes*, OCDE-CAD, Paris

<sup>8</sup> Un critère de référence ou indice se définit comme un ensemble d'indicateurs associés permettant d'établir des comparaisons constructives, précises et systématiques en matière de performance ; une norme ou grille critériée est un ensemble de critères de référence/indice ou indicateurs fournissant des informations socialement significatives en matière de performance (voir la Note n° 11, Élaboration et sélection d'indicateurs du bien-être des enfants et la Note n° 4, Logique d'évaluation).

### Points principaux

1. Les critères d'évaluation précisent les valeurs qui seront utilisées dans le cadre d'une évaluation des interventions (c.-à-d. des programmes et des politiques).
2. L'UNICEF utilise une palette de critères pour guider ses évaluations ; tous les critères ne sont pas utilisés dans chacune des évaluations, certains n'étant adaptés qu'à certains types d'évaluations ou d'interventions. Les évaluations d'impact doivent prendre en compte les éléments suivants :
  - la pertinence, l'efficacité, l'efficience, l'impact et la durabilité ;
  - l'équité, l'égalité entre les sexes et les droits de l'homme ;
  - la couverture, la coordination, la protection et la cohérence (notamment en ce qui concerne l'aide humanitaire).
3. Les critères d'évaluation « génériques » doivent être spécifiés dans les TdR d'une évaluation ou au cours des étapes initiales du processus d'évaluation, de façon à les appliquer de façon systématique et transparente, et à permettre d'aboutir à des jugements évaluatifs sur l'intervention.

## 2. QUAND UTILISER LES CRITÈRES D'ÉVALUATION ?

Toutes les évaluations d'impact conduites dans le cadre des actions de l'UNICEF doivent examiner attentivement les critères de l'OCDE-CAD et déterminer ceux qui sont pertinents pour l'évaluation en question. Ces critères reflètent les principes fondamentaux pour l'évaluation de l'aide au développement<sup>9</sup> et constituent pour la plupart des organismes d'aide au développement des normes de bonnes pratiques pour l'évaluation. Les critères d'équité, d'égalité entre les sexes et de droits de l'homme sont également exigés dans les évaluations d'impact (et autres types d'évaluation) de l'UNICEF.

Selon le type d'intervention (p. ex., fourniture d'une aide humanitaire, comme décrit plus haut) et/ou le type d'évaluation (p. ex., une évaluation de processus), il pourrait être nécessaire d'appliquer des critères supplémentaires ou de mettre l'accent sur des critères particuliers. À titre d'exemple, il ne sert à rien d'appliquer le critère d'impact de l'OCDE-CAD à une évaluation de processus, étant donné que ce type d'évaluation examine la façon dont les interventions sont mises en œuvre (notamment, les modalités de fourniture des services, la satisfaction des clients quant à ces services, les techniques de gestion employées) et non si elles ont produit les résultats escomptés.

## 3. COMMENT UTILISER LES CRITÈRES D'ÉVALUATION ?

### Associer les critères d'évaluation aux questions d'évaluation

Les TdR pour l'évaluation doivent préciser l'objectif global et la portée de l'évaluation ainsi que les critères d'évaluation génériques à prendre en compte. Ils doivent également définir les questions clés d'évaluation, bien que celles-ci puissent être élaborées ou affinées dans le cadre de la première étape du processus d'évaluation (voir la Note n° 1, Présentation de l'évaluation d'impact).

Pour que l'évaluation d'impact reste ciblée, seul un nombre limité de questions clés d'évaluation ou de questions d'évaluation de haut niveau (niveau macro) doivent être spécifiées. Par exemple :

<sup>9</sup> Organisation de coopération et de développement économiques, Comité d'aide au développement, 1991, *Principes pour l'évaluation de l'aide au développement*, OCDE-CAD, Paris. Disponible à l'adresse : [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD\(91\)208&docLanguage=Fr](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD(91)208&docLanguage=Fr)

- Question 1. Quelle était la qualité de la conception et du contenu de l'intervention ? Pertinence, équité, égalité entre les sexes, approche de la programmation fondée sur les droits de l'homme
- Question 2. Dans quelle mesure l'intervention a-t-elle été mise en œuvre et adaptée, le cas échéant ? Efficacité, efficience
- Question 3. L'intervention a-t-elle produit les résultats escomptés à court, moyen et long terme ? Dans l'affirmative, pour qui, dans quelle mesure et dans quelles circonstances ? Efficacité, impact, équité, égalité entre les sexes
- Question 4. Quels résultats imprévus (positifs et négatifs) l'intervention a-t-elle produits ? Comment se sont-ils produits ? Efficacité, impact, équité, égalité entre les sexes, approche de la programmation fondée sur les droits de l'homme
- Question 5. Quels ont été les obstacles et les facteurs favorables ayant permis de faire la différence entre une mise en œuvre de l'intervention et des résultats satisfaisants et décevants ? Pertinence, équité, égalité entre les sexes, approche de la programmation fondée sur les droits de l'homme
- Question 6. Quel intérêt les résultats ont-ils présenté pour les prestataires de services, les usagers, la communauté et/ou les organisations impliqués ? Pertinence, équité, égalité entre les sexes, approche de la programmation fondée sur les droits de l'homme
- Question 7. Dans quelle mesure l'intervention représente-t-elle la meilleure utilisation possible des ressources disponibles pour atteindre les résultats ayant la plus grande valeur ajoutée possible pour les participants et la communauté ? Efficience
- Question 8. Certains résultats positifs sont-ils susceptibles d'être durables ? Dans quelles circonstances ? Durabilité, équité, égalité entre les sexes, approche de la programmation fondée sur les droits de l'homme

Une série de questions d'évaluation plus précises (de niveau intermédiaire et inférieur) doit ensuite être formulée afin de prendre en compte chaque critère d'évaluation en détail. Toutes ces questions doivent être explicitement liées aux critères d'évaluation pour s'assurer que les critères sont intégralement couverts. Les encadrés 1 à 5 fournissent des exemples, basés sur les critères de l'OCDE-CAD, de questions de niveau intermédiaire (méso-niveau) qui sont ensuite décomposées en questions de niveau inférieur (niveau micro). Les termes niveau supérieur/macro, intermédiaire/méso et inférieur/micro n'ont pas vocation à établir une hiérarchisation en fonction de l'importance, mais plutôt d'accroître les niveaux de spécification ou de mise en œuvre opérationnelle.

### Encadré 1. Exemples de questions d'évaluation de niveaux intermédiaire et inférieur pour étudier la « pertinence »

#### **L'intervention était-elle conforme aux engagements et au plan stratégique du pays visant à accroître le bien-être des enfants ?**

- Dans quelle mesure les objectifs de l'intervention étaient-ils pertinents pour les priorités et les besoins nationaux déterminés ?

#### **La conception de l'intervention était-elle solide sur le plan technique ?**

- Les objectifs et la conception de l'intervention étaient-ils adaptés au contexte et aux besoins des bénéficiaires ?
- La logique de l'intervention était-elle claire et y avait-il des données probantes pertinentes à l'échelle locale pour relier les activités et les opérations de l'intervention avec les réalités auxquelles sont confrontés les bénéficiaires ? Une logique claire a-t-elle été fournie pour les activités de

l'intervention sélectionnées ?

- Les activités et les résultats de l'intervention étaient-ils cohérents avec le but général et la réalisation de ses objectifs ? L'étaient-ils avec les résultats et les impacts attendus ?
- Les fonds alloués ont-ils reflété de façon adéquate le niveau des besoins et l'environnement opérationnel ?

**L'intervention a-t-elle été favorable à l'égalité entre les sexes et aux autres normes relatives aux droits de l'homme ?**

- Dans quelle mesure les objectifs et processus liés à l'équité, à l'égalité entre les sexes et à l'approche de la programmation fondée sur les droits de l'homme étaient-ils intégrés à la planification de l'intervention ?
- L'intervention a-t-elle été conçue pour assurer la participation égale de l'ensemble des groupes concernés (p. ex., les hommes et les femmes, les garçons et les filles) ?
- L'intervention a-t-elle favorisé l'égalité d'accès des hommes et des femmes, des garçons et des filles aux bénéfices résultant de l'activité, et plus généralement aux ressources, aux services et aux compétences ?

**L'intervention a-t-elle été réajustée au cours de la mise en œuvre pour être mise en adéquation avec les priorités/besoins émergents et faire en sorte d'appuyer les meilleures pratiques ?**

- Dans quelle mesure l'intervention a-t-elle comporté des activités et des processus/mécanismes visant à favoriser l'efficacité de la conception et de la mise en œuvre du programme et à connaître les réactions des intéressés ?

**Encadré 2. Exemples de questions d'évaluation de niveaux intermédiaire et inférieur pour étudier « l'efficacité »**

**L'intervention a-t-elle été mise en œuvre comme prévu initialement ?**

- L'intervention a-t-elle été mise en œuvre comme prévu initialement ? Dans la négative, pour quelle raison ? Quelles mesures ont été prises pour y remédier ?
- Dans quelle mesure les facteurs contextuels ont-ils favorisé ou entravé la mise en œuvre de l'intervention ?

**Des mesures correctives ont-elles été prises lorsque le besoin s'en est fait ressentir ?**

- Un appui supplémentaire pour aider à surmonter les difficultés de mise en œuvre a-t-il été mis en évidence ou reçu ? Quelle forme ce soutien a-t-il pris, qui l'a fourni et quels en ont été les résultats ?
- Quels risques étaient associés à la réalisation des objectifs de l'intervention ? Ces risques ont-ils été correctement gérés ?

**Les résultats escomptés ont-ils été obtenus ?**

- Dans quelle mesure les objectifs de l'intervention ont-ils été atteints ? Quels ont été les principaux facteurs ayant influencé la réalisation ou non des objectifs ?

**Encadré 3. Exemples de questions d'évaluation de niveaux intermédiaire et inférieur pour étudier « l'efficacité »**

**L'intervention a-t-elle eu recours aux ressources disponibles de la façon la plus économique possible pour atteindre ses objectifs ?**

- Les objectifs ont-ils été réalisés en temps voulu ?
- L'intervention disposait-elle de suffisamment de personnel ? Celui-ci était-il compétent ? Lors de la mise en œuvre de l'intervention, les ressources et le temps ont-ils été utilisés de manière efficace pour atteindre les résultats ?
- Dans quelle mesure la coordination et la collaboration efficaces avec les interventions et les partenaires existants ont-elles été prises en compte et établies ?
- L'intervention a-t-elle été conçue et/ou modifiée au cours de la mise en œuvre afin d'optimiser les ressources financières ?
- L'intervention s'est-elle révélée aussi efficace ou plus efficace que les autres approches valables dans le même contexte ?
- Existe-t-il d'autres moyens plus économiques pour mettre en œuvre l'intervention ?

**Encadré 4. Exemples de questions d'évaluation de niveaux intermédiaire et inférieur pour étudier « l'impact »**

**L'intervention a-t-elle été suffisamment développée pour obtenir les impacts prévus (p. ex., principaux indicateurs de santé, inégalités) ?**

- Existe-t-il des preuves de changement (positif ou négatif) ? Dans l'affirmative, quels facteurs ont contribué à ce changement ? Dans la négative, pour quelle raison ?
- La vie des bénéficiaires et leur environnement ont-ils connu des changements (positifs ou négatifs) imprévus ? De quels changements s'agit-il ? Étaient-ils directement ou indirectement liés au programme ou plutôt causés par des facteurs externes ?

**Les résultats obtenus respectent-ils l'équité, l'égalité entre les sexes et les autres droits de l'homme ?**

- Dans quelle mesure l'intervention est-elle parvenue à impliquer les femmes et les hommes, les titulaires de droits et les débiteurs d'obligations, en particulier les plus vulnérables ?
- Dans quelle mesure les différents groupes, y compris les enfants et autres groupes vulnérables, ont-ils bénéficié de l'intervention, sur différents plans ?
- Dans quelle mesure les capacités des différents groupes, y compris des enfants et autres groupes vulnérables, à interagir différemment avec d'autres groupes au cours de l'intervention, ont-elles été renforcées ?
- Comment les coûts et les inconvénients de la participation à l'intervention ont-ils été répartis entre les différents groupes de participants ? Qu'en est-il des avantages comparatifs, le cas échéant, tirés de leur participation ?



### Le contexte opérationnel a-t-il exercé une influence sur l'intervention ou inversement ?

- Dans quelle mesure le contexte local et national a-t-il participé au changement ?
- L'intervention a-t-elle eu un impact sur le contexte opérationnel plus large œuvrant au bien-être des enfants dans le pays ? Dans l'affirmative, quelle forme cela a-t-il pris ?

### Encadré 5. Exemples de questions d'évaluation de niveaux intermédiaire et inférieur pour étudier la « durabilité »

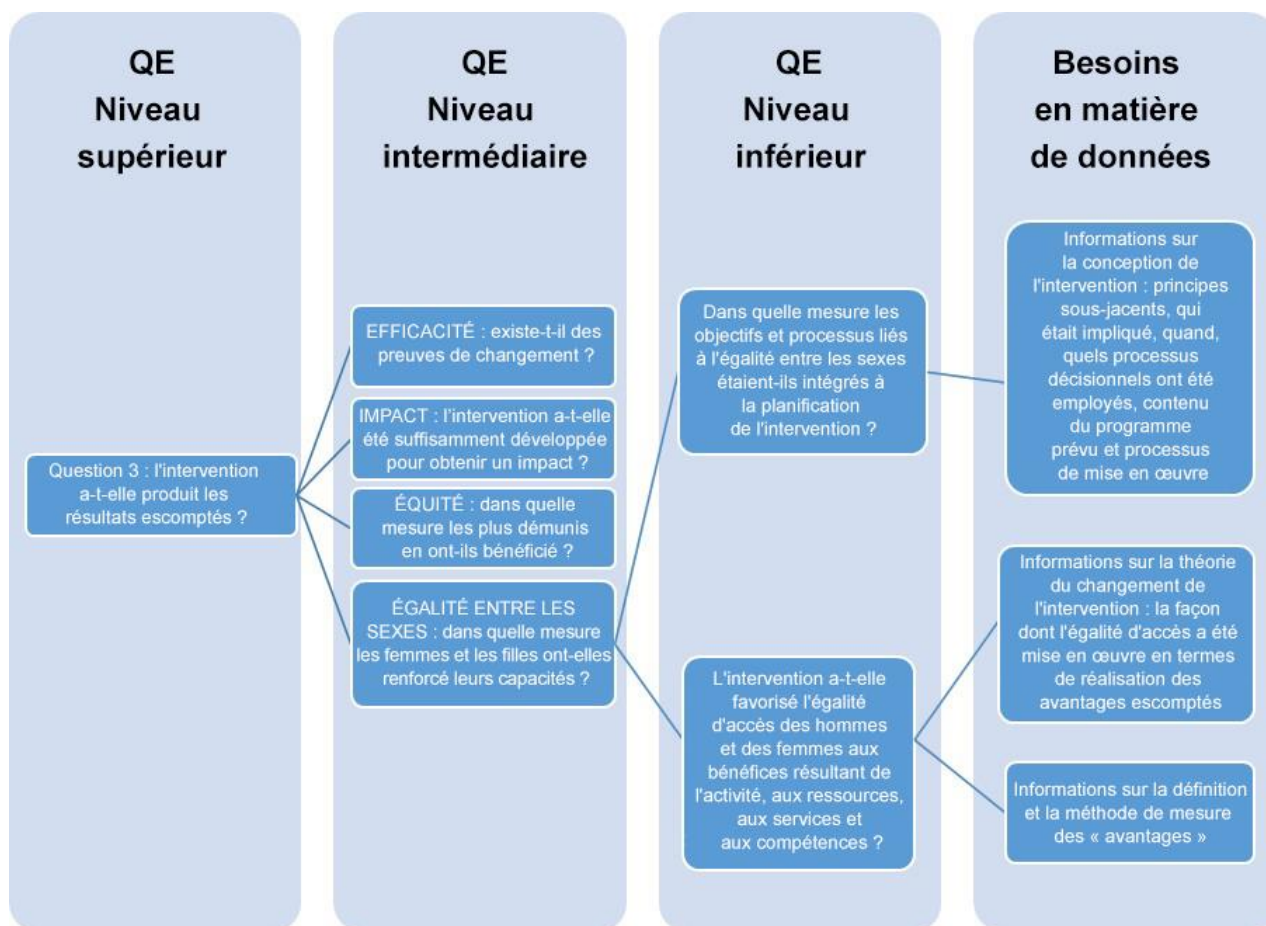
#### Les résultats ont-ils été obtenus de manière durable ?

- Dans quelle mesure les activités et les avantages de l'intervention perdurent-ils lorsque le financement externe prend fin ?
- Certains domaines d'intervention sont-ils à l'évidence non durables ? Quels enseignements peut-on en tirer ?
- Dans quelle mesure les bénéficiaires de l'intervention et/ou les parties prenantes des pays partenaires se la sont-ils appropriée pour perpétuer ses résultats au moment où le financement externe prend fin, en ont-ils la capacité et disposent-ils des ressources pour le faire ?
- L'intervention a-t-elle contribué au renforcement des capacités des organisations locales afin que celles-ci continuent à fournir des services de qualité aux enfants ? Dans l'affirmative, quelle forme cela a-t-il pris ?
- Quels ont été les principaux facteurs ayant influencé la (non-)durabilité de l'intervention ?

Un plan d'évaluation doit être élaboré en réponse aux TdR de l'évaluation. Il doit décrire les questions d'évaluation principales et les questions détaillées liées aux critères d'évaluation génériques et spécifiques (ou la façon dont ces derniers seront définis dans le cadre du processus d'évaluation) ; les données nécessaires pour répondre à chacune des questions d'évaluation, ainsi que la méthode de recherche et d'évaluation associée (voir la Note n° 6, Présentation des stratégies d'attribution causale) ; et les méthodes de collecte/regroupement de données et d'analyse (voir la Note n° 10, Présentation des méthodes de collecte et d'analyse de données). Le plan d'évaluation doit également comprendre les responsabilités (quelles seront les attributions de chacun), le calendrier et les livrables afin qu'il soit également utilisé par les personnes mandatées à l'évaluation et par ses gestionnaires afin de maintenir le processus d'évaluation sur la bonne voie.

La figure 1 résume la façon dont les critères d'évaluation sont liés aux questions d'évaluation (aux niveaux inférieur, intermédiaire et supérieur), ainsi qu'aux besoins en matière de données dans le cas de la question 3 (voir plus haut) : l'intervention a-t-elle produit les résultats escomptés à court, moyen et long terme ? Dans l'affirmative, pour qui, dans quelle mesure et dans quelles circonstances ?

Figure 1. Exemple illustrant l'établissement de liens entre une question d'évaluation et un critère d'évaluation d'une part, et des besoins de données spécifiques d'autre part



Remarque : QE = question d'évaluation

### Utilisation de critères de référence ou de normes spécifiques pour apprécier la performance de l'intervention

De bonnes questions d'évaluation ne portent pas uniquement sur les résultats obtenus (questions descriptives sur la situation initiale et son évolution), mais également sur la *qualité* des résultats (questions évaluatives portant sur la conclusion générale, à savoir si le programme ou la politique a atteint les objectifs escomptés).

Les TdR de l'évaluation peuvent inclure des critères de référence ou des normes spécifiques pour apprécier l'intervention ou, lorsque ces derniers n'existent pas encore, la façon de les définir collectivement au cours du processus d'évaluation. À titre d'exemple, si l'efficacité est définie en fonction de la mesure dans laquelle les objectifs ont été atteints et que l'un des objectifs est « allongement de la période de scolarisation des enfants », le critère spécifique pour apprécier l'intérêt de l'intervention pourrait être la rétention scolaire (rester à l'école jusqu'aux dernières années) et la fréquentation scolaire (p. ex., tous les enfants assistent à 90 % des cours, ou 90 % des élèves assistent à l'ensemble des cours).

Il existe des critères de référence et/ou des normes propres à certains domaines et/ou contextes liés à l'enfant. Ils sont généralement fondés sur des données probantes et élaborés en concertation avec les parties prenantes concernées. Leur but principal est de remplir la fonction d'objectifs spécifiques (réalistes plutôt qu'ambitieux), pour suivre (à l'échelle locale, nationale et/ou mondiale) les progrès et évaluer les réalisations. Lorsqu'ils sont correctement définis et mis en œuvre, ils peuvent être utilisés pour comparer et

tirer des enseignements des performances des programmes/politiques, des différents contextes et/ou de zones géographiques spécifiques. À titre d'exemple, des critères de référence et/ou des normes ont été définis pour :

- [les objectifs de couverture vaccinale de l'OMS](#) : leur raison d'être réside dans le fait que la couverture minimale est essentielle pour une prévention efficace des épidémies ;
- [les systèmes nationaux de protection de l'enfance de l'UNICEF](#) (voir l'encadré 6) : leur raison d'être réside dans le fait que les systèmes nationaux de protection de l'enfance sont reconnus comme étant les moyens les plus efficaces et durables de protéger les enfants contre toute forme de maltraitance. L'absence d'un tel système conduit souvent à une couverture incomplète, et à la prise en charge des violations individuelles en tant que problèmes distincts et autonomes<sup>10</sup>.

Les critères d'évaluation ne sont pas seulement utiles pour évaluer les interventions, mais ils doivent également être explicites dans la [théorie du changement](#) (voir la Note n° 2, Théorie du changement) et rendus opérationnels dans le contenu de l'intervention et dans la façon dont celle-ci est mise en œuvre. L'encadré 6 présente un exemple tiré de l'UNICEF de normes et critères de référence utilisés pour évaluer le secteur de la protection de l'enfance.

### Encadré 6. Normes et critères de référence utilisés dans la prestation et l'évaluation des services de protection de l'enfance

#### But

Pour soutenir le personnel œuvrant dans la prestation de services de protection de l'enfance, il est essentiel que les pays mettent en place des systèmes complets basés sur les obligations légales respectives, une formation continue sur ce sujet pour maintenir leurs connaissances à jour et un système d'inspection régulière des niveaux de compétences.

#### Critères de référence du programme

Le Bureau régional pour l'Asie orientale et le Pacifique de l'UNICEF (UNICEF EAPRO) a proposé les **critères de référence** suivants relatifs à l'enseignement professionnel et au système de développement continu :

- Un programme de diplômes universitaires en action sociale, dont les programmes d'études comprennent des matières relatives aux services sociaux, aux questions de développement, aux sujets liés à la protection et à la prévention, et aux interventions thérapeutiques.
- Un programme de qualification professionnelle en action sociale ou en développement de l'enfant, disposant d'un programme d'études validé par les autorités compétentes.
- Un système d'accréditation des compétences en action sociale, s'appuyant sur des tests de compétence dans le cadre de programmes de formation adaptés.
- Une formation sur la gestion de la maltraitance, de la violence et de l'exploitation destinée aux éducateurs (p. ex., les enseignants), aux professionnels de santé et/ou aux autres professionnels travaillant avec des enfants.
- Un programme de formation spécifique sur les enfants et la justice des mineurs destiné au personnel des ministères occupant des postes de direction aux affaires intérieures ou à la justice.

#### Jugements évaluatifs

<sup>10</sup> Fonds des Nations Unies pour l'enfance, 2012, *Measuring and Monitoring Child Protection Systems: Proposed Regional Core Indicators for East Asia and the Pacific*, Strengthening Child Protection Systems No. 3, UNICEF EAPRO, Bangkok, avant-propos, p. vii. Disponible à l'adresse : [http://www.unicef.org/eapro/Measuring\\_and\\_monitoring.pdf](http://www.unicef.org/eapro/Measuring_and_monitoring.pdf)

Des **normes** ont ensuite été définies pour évaluer la situation d'un système de formation professionnelle national :

- Norme A : le système éducatif comprend **l'ensemble** des éléments figurant dans les critères de référence.
- Norme B : le système éducatif comprend **au moins trois** éléments figurant dans les critères de référence.
- Norme C : le système éducatif comprend **un ou deux** éléments figurant dans les critères de référence.
- Norme D : le système éducatif ne comprend **aucun** des éléments figurant dans les critères de référence.

Ces normes pourraient par exemple être utilisées pour apprécier l'intérêt du programme de développement professionnel d'un pays en formulant un système de notation. Par exemple :

- **Excellent** – exemple clair de performances exemplaires ou de meilleure pratique dans ce domaine ; absence de points faibles ; répond à la norme A.
- **Très bien** – performance très bonne ou excellente sur pratiquement tous les aspects ; globalement solide, mais pas exemplaire ; absence de points faibles ayant une réelle importance ; répond à la norme B et, par ailleurs, à l'ensemble des exigences de qualité pour la formation.
- **Inacceptable** – points faibles majeurs dans tous les domaines du programme ; répond à la norme C ou D.

Source : Fonds des Nations Unies pour l'enfance, 2012, *Measuring and Monitoring Child Protection Systems: Proposed Regional Core Indicators for East Asia and the Pacific*, Strengthening Child Protection Systems No. 3, UNICEF EAPRO, Bangkok, avant-propos, p. vii. Disponible à l'adresse : [http://www.unicef.org/eapro/Measuring\\_and\\_monitoring.pdf](http://www.unicef.org/eapro/Measuring_and_monitoring.pdf).

La réflexion relative à la façon de définir – et donc d'apprécier – la réussite d'un programme ou d'une politique doit être menée en amont. Ainsi, la collecte de données peut permettre d'obtenir l'ensemble des données probantes nécessaires pour porter des jugements appropriés sur la performance. Définir la « réussite » avant de collecter les données constitue par ailleurs une stratégie plus transparente et objective. Faire participer les parties prenantes concernées à la définition de la « réussite », pour une intervention donnée et dans un contexte donné, améliore l'acceptation des résultats de l'évaluation (en réduisant les malentendus et les désaccords relatifs à l'interprétation par la suite) et incite à prendre les mesures nécessaires en ce qui concerne le programme. La Note n° 4, Logique d'évaluation, définit une approche par étapes pour décider (1) des aspects pour lesquels l'intervention doit obtenir de bons résultats (c.-à-d., établir des critères spécifiques définissant « l'équité »), et (2) des performances de l'intervention sur ces aspects (c.-à-d., élaborer une norme ou une [grille critériée](#) ou en utiliser une déjà existante).

## 4. QUESTIONS ÉTHIQUES ET LIMITES PRATIQUES

### Questions éthiques

Les évaluations de l'UNICEF doivent suivre les *Ethical Guidelines for Evaluation*<sup>11</sup> (directives éthiques pour l'évaluation) définies par le Groupe des Nations Unies pour l'évaluation (GNUE). Ces directives contiennent un code de conduite pour les évaluateurs ainsi que des obligations d'ordre éthique pour les

<sup>11</sup> Groupe des Nations Unies pour l'évaluation, 2007, *Ethical Guidelines for Evaluation*, GNUE, New York.

personnes participant à l'évaluation (voir également la Note n° 1, Présentation de l'évaluation d'impact). Ces obligations concernent notamment : le respect de la dignité et de la diversité, les droits, la confidentialité et le principe « ne pas nuire ». Ces obligations doivent être respectées de façon explicite dans le modèle d'évaluation et tout au long du processus. Elles ont pour but de garantir le respect et la protection des droits et du bien-être des personnes et de leurs communautés, conformément à la Déclaration universelle des droits de l'homme et aux autres conventions relatives aux droits de l'homme.

En octobre 2013, le Centre de recherche de l'UNICEF a créé [Ethical Research Involving Children](#) (Recherche éthique impliquant des enfants), une ressource en ligne<sup>12</sup> qui réunit des experts qui examinent certaines questions éthiques clés concernant les enfants et la manière de traiter ces questions dans différents contextes de recherche (et d'évaluation).

Outre les questions éthiques générales susmentionnées, on peut citer des questions particulières liées aux choix effectués en matière de critères d'évaluation, notamment :

- Qui doit participer à la définition de ce qui constitue une réussite ? Il convient également de décider qui a le dernier mot lorsque des priorités/points de vue divergents émergent ou lorsque des données probantes contradictoires apparaissent.
- Comment définir la réussite ? La définir uniquement en termes « d'effets moyens » ou « d'indicateurs sociaux globaux » pourrait passer sous silence l'expérience de certains sous-groupes (tels que les garçons ou les filles, les plus vulnérables ou les plus pauvres).
- Comment les résultats escomptés ont-ils été obtenus ? Par exemple, est-il contraire à l'éthique d'accroître la couverture vaccinale par le biais d'une participation prescrite par l'autorité désignée afin de lutter contre une épidémie ?

### Limites pratiques

Les limites pratiques fréquemment rencontrées dans la prise en compte des critères d'évaluation sont les suivantes :

- La difficulté de passer d'une formulation générique des critères d'évaluation (comme le critère de l'OCDE-CAD de « durabilité ») à des critères plus spécifiques prenant en compte le point de vue de différentes parties prenantes sur les aspects importants et les seuils que l'intervention doit dépasser.
- Les incohérences entre le calendrier de l'évaluation d'impact et l'apparition des impacts concernés ou des données relatives à la durabilité. Mesurer des « indicateurs indirects » (c.-à-d., ce qu'il est possible de mesurer plutôt que ce qui devrait être mesuré dans l'idéal) pourrait compliquer davantage les définitions et les négociations à propos de la réussite.
- Les informations relatives à la qualité de certaines composantes ou de certains services mis en œuvre dans le cadre d'un programme à grande échelle sont généralement limitées. Elles sont souvent compliquées à remettre en place ou à comprendre *a posteriori*, et donc difficiles à bien cerner au moment de préciser les critères d'évaluation.

Si les ressources et le calendrier de l'évaluation le permettent, certaines de ces limites peuvent être corrigées (p. ex., en prolongeant la période de l'évaluation ou sa portée), mais il est probable que les conclusions devront être tirées de données imparfaites. Par conséquent, il est important de préciser ouvertement les limites de l'évaluation et de décrire comment elles ont pu affecter les résultats, les conclusions et les recommandations sur un plan général.

---

<sup>12</sup> Les nouvelles ressources comprennent : la Charte internationale pour la recherche éthique impliquant des enfants ; un site Internet, [www.childethics.com](http://www.childethics.com) (en anglais), qui fournit un référentiel riche d'informations basées sur des données probantes, des ressources et des liens vers des articles de revues afin d'orienter et d'améliorer la recherche impliquant les enfants, ainsi qu'une plateforme pour faciliter une plus ample réflexion et le dialogue critique ; et un recueil de questions éthiques et de défis, y compris une collection de plus de 20 études de cas ainsi que des questions structurées pour orienter la recherche éthique impliquant des enfants (page Internet intitulée « [Getting Started](#) »).

## 5. QUELLES SONT LES AUTRES MÉTHODES QUI FONCTIONNENT BIEN AVEC CELLE-CI ?

Toutes les évaluations consistent à poser des questions évaluatives et à y répondre – et non uniquement à mesurer et à décrire des résultats. De ce fait, tous les résultats doivent être interprétés dans un cadre d'évaluation. Cela implique non seulement d'énoncer les résultats, mais également de se prononcer sur leur qualité. Pour ce faire, il est nécessaire de mettre en place une logique d'évaluation, qui est une condition *sine qua non* à l'ensemble des évaluations, indépendamment des méthodes ou des approches utilisées (voir la Note n° 4, Logique d'évaluation).

Dans le contexte de l'UNICEF, de nombreuses évaluations d'impact reposeront sur l'utilisation de sources de données existantes. Il importe donc de procéder également à une évaluation fondée sur des normes<sup>13</sup> de la qualité du système de suivi et d'évaluation du programme et de la situation stratégique globale. L'utilisation de données dont la qualité n'est pas connue pourrait compromettre la validité des résultats de l'évaluation et donner lieu à des conclusions erronées. Les évaluateurs doivent utiliser et collecter les meilleures données disponibles et s'assurer que les conséquences de tout problème de qualité des données soient minimisées, dans la mesure du possible, et spécifiquement mentionnées dans le rapport d'évaluation.

## 6. PRÉSENTATION ET ANALYSE DES RÉSULTATS

Le rapport d'évaluation doit être structuré de manière à refléter le but et les questions clés de l'évaluation (voir également la Note n° 1, Présentation de l'évaluation d'impact).

Dès le début, les données probantes permettant de répondre aux questions détaillées liées aux critères de pertinence, d'efficacité, d'efficience, d'impact et de durabilité de l'OCDE-CAD, ainsi que les considérations d'équité, d'égalité entre les sexes et de droits de l'homme, doivent être présentées brièvement, mais avec suffisamment de détails pour étayer les conclusions et les recommandations. Présenter les données probantes de cette manière aidera l'[équipe de gestion de l'évaluation](#) à s'assurer que les critères définis dans les TdR soient pleinement pris en compte.

Les grilles critériées d'évaluation spécifiques doivent être utilisées pour « interpréter » les données probantes et déterminer les considérations présentant un caractère absolument essentiel ou impérieux. Les données probantes sur divers aspects doivent donc être synthétisées pour répondre aux questions d'évaluation de haut niveau. Des recommandations plus spécifiques sur la façon de rédiger un rapport rigoureux sur la synthèse d'évaluation sont fournies dans la Note n° 4, Logique d'évaluation (p. ex., en utilisant clairement des termes spécifiques à l'évaluation pour présenter les résultats, au lieu de termes neutres qui ne font que décrire les résultats).

<sup>13</sup> Des outils basés sur des normes convenues à l'échelle mondiale pour évaluer un système de suivi et d'évaluation (aux échelles nationales, infranationales et de la prestation de services) sont disponibles. Par exemple, Programme commun des Nations Unies sur le VIH/sida, 2009, *12 Components Monitoring and Evaluation System Strengthening Tool*, ONUSIDA, Genève. Disponible à l'adresse : [http://www.unaids.org/en/media/unaids/contentassets/documents/document/2010/2\\_MERG\\_Strengthening\\_Tool\\_12\\_Component\\_s\\_ME\\_System.pdf](http://www.unaids.org/en/media/unaids/contentassets/documents/document/2010/2_MERG_Strengthening_Tool_12_Component_s_ME_System.pdf)

## 7. EXEMPLES DE BONNES PRATIQUES

Certains exemples précis de bonnes pratiques sont présentés dans l'encadré 7 ci-dessous.

### Encadré 7. Exemple tiré d'une évaluation d'impact permanente de programmes conjoints sur l'égalité entre les sexes du système des Nations Unies

#### Bonnes pratiques en matière de termes de référence

Sept questions clés d'évaluation ont été incluses. Elles étaient directement liées aux priorités stratégiques et aux questions émergentes définies dans l'analyse des programmes, par exemple :

- Dans quelle mesure les programmes communs sur l'égalité des sexes ont-ils été conceptualisés, planifiés et élaborés ensemble pour répondre aux engagements pris à l'échelle nationale, régionale et internationale en matière d'égalité des sexes et d'autonomisation des femmes ? Pour garantir une cohérence et exploiter les avantages comparatifs des organismes des Nations Unies participants ? Et pour intégrer une approche de la programmation fondée sur les droits de l'homme ?
- Quelles sont les principales contributions et la valeur ajoutée des résultats sur le court et long terme, escomptés et imprévus, positifs et négatifs, relatifs à l'égalité des sexes et à l'autonomisation des femmes, obtenus à ce jour par les programmes communs sur l'égalité des sexes à l'échelle nationale ?
- Quelle influence exercent le contexte et les circonstances spécifiques du pays (pays en conflit/sortant d'un conflit, pays fragile, pays adhérant à l'initiative « Unis dans l'action ») sur l'obtention des résultats du programme commun sur l'égalité des sexes et l'efficacité opérationnelle ?

Les critères d'évaluation pour évaluer les programmes communs sur l'égalité des sexes ont été formulés et définis dans le contexte du programme (au lieu d'être uniquement fondés sur la description générique de l'OCDE-CAD). Par exemple :

- La pertinence/cohérence des processus de planification, de conception et de mise en œuvre des programmes communs sur l'égalité des sexes par rapport : aux engagements, aux politiques et aux priorités à l'échelle nationale, régionale et internationale ; aux principes d'efficacité de l'aide ; aux missions des Nations Unies et au Plan-cadre des Nations Unies pour l'aide au développement ; aux politiques, missions et avantages comparatifs individuels des différents organismes quant à leur réactivité et à leur prise en compte des besoins des pays en matière d'égalité des sexes et d'autonomisation des femmes.
- L'efficacité et l'impact des programmes communs sur l'égalité des sexes quant à leur capacité à atteindre leurs objectifs énoncés en matière d'égalité des sexes et d'autonomisation des femmes et tout effet à long terme escompté ou imprévu à l'échelle nationale, ainsi que leur capacité à intégrer l'égalité des sexes dans le système des Nations Unies, notamment en ce qui concerne l'utilisation d'approches innovantes.
- La durabilité des résultats des programmes communs sur l'égalité des sexes au regard du niveau d'appropriation généré dans le pays, de l'efficacité des partenariats établis et du renforcement des capacités nationales grâce aux processus desdits programmes.

Des critères supplémentaires ont été mis en place, comme la participation et l'inclusion de titulaires de droits et de débiteurs d'obligations à l'échelle nationale – en particulier les plus marginalisés – dans les processus de programmes communs sur l'égalité des sexes.

La période de la planification de l'évaluation a également permis d'introduire de nouveaux critères d'évaluation avant de convenir de la portée définitive de l'évaluation.

Outre les questions clés d'évaluation, des questions similaires plus détaillées ont été incluses dans une « matrice des questions d'évaluation » en annexe des TdR.

Source : voir les termes de référence relatifs aux évaluations et les rapports thématiques du Fonds des Nations Unies pour la population, « Joint Evaluation of Joint Gender Programmes in the UN System », UNFPA, <http://www.unfpa.org/admin-resource/joint-evaluation-joint-gender-programmes-un-system>

## 8. EXEMPLES DE DIFFICULTÉS

Certaines lacunes communes dans la prise en compte des critères d'évaluation ont été identifiées, par exemple, dans l'examen de la qualité de l'ensemble des rapports d'évaluation de 2012<sup>14</sup>.

- « Les questions et les critères sous-jacents guidant l'évaluation ne sont pas précisés. »
- « L'intégration des aspects relatifs à l'approche de la programmation fondée sur les droits de l'homme, l'égalité entre les sexes et l'équité reste globalement insuffisante dans les évaluations (à savoir, la proportion de rapports intégrant avec succès ces éléments n'a pas dépassé 50 % ; 46 % des rapports ont pris en compte l'égalité entre les sexes, 44 % l'approche de la programmation fondée sur les droits de l'homme et seuls 41 % des rapports réservent une plus grande place à l'équité). »

Les remarques suivantes ont également été formulées : « Les rapports d'évaluation ont tendance à refléter ce qu'il a été demandé aux évaluateurs de prendre en compte, notamment les méthodes suggérées, les critères d'évaluation, les sources de données et les questions précisées dans les [termes de référence]. Ainsi, même lorsque les questions d'égalité entre les sexes, de droits de l'homme et d'équité sont essentielles aux valeurs de l'Organisme, leur prise en compte dans les rapports d'évaluation a tendance à être dictée par les [termes de référence]. »

De ce fait, il est important de fournir aux évaluateurs des orientations, des TdR et des commentaires cohérents et explicites, au cours du processus d'évaluation.

## 9. PRINCIPALES LECTURES ET LIENS UTILES

BetterEvaluation, « Define ethical and quality evaluation standard », page Internet, [http://betterevaluation.org/plan/manage\\_evaluation/ethical\\_evaluation](http://betterevaluation.org/plan/manage_evaluation/ethical_evaluation)

BetterEvaluation, « Establish Decision Making Processes », page Internet, [http://betterevaluation.org/plan/manage/who\\_controls](http://betterevaluation.org/plan/manage/who_controls)

BetterEvaluation, « Evaluation Plan », page Internet, [http://betterevaluation.org/evaluation-options/evaluation\\_planning\\_templates](http://betterevaluation.org/evaluation-options/evaluation_planning_templates)

BetterEvaluation, « Specify the Key Evaluation Questions », page Internet, [http://betterevaluation.org/plan/engage\\_frame/decide\\_evaluation\\_questions](http://betterevaluation.org/plan/engage_frame/decide_evaluation_questions)

Centre de recherche de l'UNICEF — Innocenti, Childwatch International Research Network, Centre for Children and Young People de la Southern Cross University (Australie) et Children's Issues Centre de la

<sup>14</sup> Au total, 79 rapports d'évaluation ont été examinés. Les citations sont tirées du résumé (p. iii) de Universal Management Group, 2013, « Geros – Global Meta-Evaluation Report 2012 », rapport final, Universal Management Group. Disponible à l'adresse : [http://www.unicef.org/evaldatabase/files/2013\\_GEROS\\_MetaEvaluation\\_finalReport.pdf](http://www.unicef.org/evaldatabase/files/2013_GEROS_MetaEvaluation_finalReport.pdf)



University of Otago (Nouvelle-Zélande), 2013, site Internet « Ethical Research Involving Children », [www.childethics.com](http://www.childethics.com)

Groupe des Nations Unies pour l'évaluation, 2007, *Ethical Guidelines for Evaluation*, GNUE, New York. Disponible à l'adresse : <http://www.unevaluation.org/ethicalguidelines>

Organisation de coopération et de développement économiques, « Evaluation of development programmes, DAC Criteria for Evaluating Development Assistance », page Internet, OCDE, <http://www.oecd.org/fr/cad/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>

Organisation de coopération et de développement économiques, Comité d'aide au développement, 1991, *Principes pour l'évaluation de l'aide au développement*, OCDE-CAD, Paris. Disponible à l'adresse : [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD\(91\)208&docLanguage=Fr](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD(91)208&docLanguage=Fr)

## GLOSSAIRE

<b><u>Équipe de gestion de l'évaluation</u></b>	<i>Équipe chargée de gérer une évaluation. À l'UNICEF, ses missions consistent à : sélectionner et superviser le travail des évaluateurs (généralement externes) ; examiner les TdR, le rapport initial et le projet de rapport final ; apporter des conseils et orientations techniques aux évaluateurs ; gérer les aspects opérationnels de l'évaluation (le budget, les visites sur le terrain et les échanges avec les contreparties de l'UNICEF concernés) ; et procéder à l'assurance qualité. Dans une évaluation conjointe, cette équipe comprend des membres du personnel de l'UNICEF et des représentants de chaque organisation partenaire. Terme associé : comité de pilotage de l'évaluation.</i>
<b><u>Grille critériée</u></b>	<i>Outil d'évaluation basé sur un ensemble de critères et un guide de scores/notes préétablis par le ou les évaluateurs. Les grilles critériées peuvent servir à évaluer des présentations, des projets, des portefeuilles, etc.</i>
<b><u>Questions clés d'évaluation</u></b>	<i>Questions d'évaluation de haut niveau (niveau supérieur) relatives à la performance générale, auxquelles l'évaluation devrait chercher à répondre. Les questions clés d'évaluation sont obtenues à partir du but de l'évaluation.</i>
<b><u>Termes de référence (TdR)</u></b>	<i>Déclaration relative au contexte, aux objectifs, aux destinataires, aux questions clés d'évaluation, à la méthodologie, aux rôles et responsabilités, au calendrier, aux livrables, aux normes de qualité, aux qualifications de l'équipe d'évaluation ainsi qu'à d'autres questions pertinentes et qui précise les fondements sous-tendant le contrat passé entre l'UNICEF et les évaluateurs.</i>
<b><u>Théorie du changement</u></b>	<i>Explique comment les activités s'entendent comme produisant un ensemble de résultats qui contribuent à la réalisation des impacts finaux prévus. Elle peut être élaborée pour tous les niveaux d'intervention : événement, projet, programme, politique, stratégie ou organisation.</i>