

Présentation de l'évaluation d'impact

Patricia Rogers

CENTRE DE RECHERCHE INNOCENTI 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.

Cette publication ne reflète pas les politiques ou points de vue de l'UNICEF sur certains sujets. 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.

NOTES MÉTHODOLOGIQUES DU CENTRE DE RECHERCHE INNOCENTI

Les notes méthodologiques du Centre d'Investigations d'UNICEF ont le but de partager des pratiques, des méthodes et des dessins 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

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:

Rogers, P. (2014). Présentation de l'évaluation d'impact, *Note méthodologique n°1*, Centre de recherche Innocenti, Florence.

Remerciements : Ce dossier 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: Simon Hearn, Jessica Sinclair Taylor, Howard White

Pour leur révision: Nikola Balvin, Samuel Bickel, Christian Salazar, David Stewart

© Fonds des Nations Unies pour l'enfance (UNICEF), septembre 2014
Centre de recherche Innocenti de l'UNICEF

Piazza SS. Annunziata, 12
50122 Florencia (Italia)
Tel: (+39) 055 20 330
Fax: (+39) 055 2033 220
florence@unicef.org
www.unicef-irc.org

1. QU'EST-CE QUE L'ÉVALUATION D'IMPACT – ET POURQUOI EST-CE IMPORTANT?

Les évaluations d'impact fournissent des informations sur les effets induits par une intervention. Il est possible de réaliser l'évaluation d'impact d'un programme, d'une politique ou d'un travail en amont, par exemple le renforcement des capacités, le plaidoyer politique et la mise en place d'un environnement favorable. Cela va au-delà d'une simple étude des buts et objectifs pour examiner également les impacts inattendus. Le Comité d'aide au développement de l'Organisation de coopération et de développements économiques (OCDE-CAD) définit les impacts comme étant des « effets à long terme, positifs et négatifs, primaires et secondaires, induits par une action de développement, directement ou non, intentionnellement ou non »¹.

Une évaluation d'impact peut être formative (pour améliorer ou réorienter une politique ou un programme) ou sommative (pour éclairer les décisions quant à l'intérêt de poursuivre, interrompre, reproduire, étendre ou non un programme ou une politique). Elle peut être utilisée par l'UNICEF et ses partenaires au niveau mondial, régional ou national pour éclairer les décisions ainsi qu'à des fins de plaidoyer et de conseils.

La présente note donne un aperçu des différents éléments de l'évaluation d'impact et des diverses options qui s'offrent aux responsables de programme de l'UNICEF pour chacun de ces éléments, et précise les étapes nécessaires pour la planification et la gestion d'une évaluation d'impact. Ces étapes sont les suivantes, bien que leur ordre et contenu puissent varier:

- décision de procéder à une évaluation d'impact ;
- détermination des dispositifs en matière de gouvernance et de gestion ;
- préparation de la mise en œuvre d'une évaluation d'impact ;
- élaboration des [termes de référence \(TdR\)](#) pour l'évaluation ;
- constitution de l'équipe d'évaluation ;
- supervision de l'évaluation, y compris la publication de rapports d'évaluation ;
- suivi de l'évaluation.

Cette note fournit également des conseils visant à répondre aux questions d'éthique et garantir la qualité au cours d'une évaluation d'impact.

Ce document constitue une introduction à une série de notes méthodologiques de l'UNICEF qui fournissent des indications plus détaillées sur les fondements essentiels d'une évaluation d'impact (théorie du changement, critères d'évaluation et logique d'évaluation), différentes approches et méthodes de collecte et d'analyse des données, ainsi que diverses options en termes d'approches participatives.

2. DÉCISION DE PROCÉDER À UNE ÉVALUATION D'IMPACT

Il est important que l'évaluation d'impact soit abordée dans le cadre d'un Plan intégré de suivi, évaluation et recherche (IMEP) conçu pour générer des données probantes en vue d'éclairer la prise de décisions. Ceci garantira que les données issues d'autres composants de suivi et d'évaluation, tels

¹ Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats, CAD/OCDE, OCDE, Paris, 2010. Voir <http://www.oecd.org/development/peer-reviews/2754804.pdf>.

que le [suivi des performances](#) et [l'évaluation de processus](#), puissent être utilisées en fonction des besoins. Ceci garantira en outre la prise en compte précoce de la planification d'une évaluation d'impact, permettant de ce fait la collecte de données de référence et, le cas échéant, la création d'un [groupe de contrôle](#) ou [groupe témoin](#), ou le recours à d'autres stratégies pour l'étude des [liens de causalité](#).

Une évaluation d'impact doit être effectuée uniquement lorsqu'il est possible de déterminer clairement son utilisation prévue et lorsqu'elle permet vraisemblablement d'aboutir à des conclusions utiles, en tenant compte de la disponibilité des ressources et du calendrier des décisions concernant le programme ou la politique en question. Il peut s'avérer nécessaire de réaliser en premier lieu une [étude d'évaluabilité](#)² formelle afin d'apprécier ces aspects.

On effectue des évaluations d'impact formatives pour éclairer les décisions relatives à la modification d'un programme ou d'une politique. Bien que de nombreuses évaluations formatives soient axées sur les processus, les évaluations d'impact peuvent être utilisées à des fins de formation dans le cadre d'intervention en cours. Les conclusions d'une évaluation d'impact peuvent par exemple servir à améliorer la mise en œuvre d'un programme pour le prochain groupe de participants.

On effectue des évaluations d'impact sommatives pour éclairer les décisions quant à l'intérêt de poursuivre, d'interrompre, de reproduire, d'étendre ou non un programme ou une politique. Idéalement, une évaluation d'impact sommative permet non seulement de constater « ce qui fonctionne », mais elle fournit aussi des informations sur ce qui est nécessaire pour que l'intervention soit efficace pour différents groupes dans différents contextes, informations qui peuvent par la suite être utilisées pour éclairer les décisions.

Tableau 1. Quand est-il approprié de mener une évaluation d'impact?

Problématique	Une évaluation d'impact est appropriée quand...	Une évaluation d'impact peut ne pas être appropriée quand...
Utilisations prévues et calendrier	Il est possible d'utiliser les constatations afin d'éclairer les décisions concernant les programmes ou politiques à venir.	Il n'existe pas d'utilisations ou d'utilisateurs prévus clairement définis – les décisions ont par exemple déjà été prises sur la base de données probantes, crédibles existantes, ou devront l'être avant qu'une évaluation d'impact fiable puisse être réalisée.
Objet principal	Il est nécessaire de comprendre les impacts produits.	La priorité absolue à ce stade est de comprendre et d'améliorer la qualité de la mise en œuvre.
Ressources	Il existe assez de ressources, notamment des données de bonne qualité, pour réaliser une évaluation d'impact suffisamment complète et précise, et l'on dispose de temps et de fonds pour en collecter davantage.	Les données existantes ne sont pas adaptées et les ressources sont insuffisantes pour combler les lacunes.

² Pour de plus amples informations, consulter la page Internet de BetterEvaluation « Evaluability Assessment », http://betterevaluation.org/themes/evaluability_assessment (en anglais).

Pertinence	Le lien avec les stratégies et priorités nationales et de l'UNICEF est clairement établi.	Le lien avec les stratégies et priorités nationales et de l'UNICEF est vague.
------------	---	---

Le processus de hiérarchisation des interventions devant faire l'objet d'une évaluation d'impact doit impliquer diverses parties prenantes, dont des représentants gouvernementaux et de la société civile, des responsables et hauts fonctionnaires du système des Nations Unies ainsi que des représentants d'autres organisations partenaires. Ce processus doit tenir compte de la pertinence de l'évaluation quant aux stratégies nationales et des Nations Unies, de son utilité potentielle et de l'engagement qu'ont pris les hauts responsables d'y avoir recours, de sa possible utilisation à des fins de plaidoyer en faveur de politiques fondées sur des données probantes, ainsi que des exigences en matière de reddition de comptes³.

Il est également primordial de tenir compte du moment opportun pour effectuer une évaluation d'impact. Les informations tirées d'évaluations d'impact effectuées tardivement seront disponibles trop tard pour permettre d'éclairer les décisions. C'est pourquoi certains rapports présentés comme des « évaluations d'impact » exposent uniquement les effets directs intermédiaires évidents au cours de l'évaluation plutôt que les impacts à long terme de l'intervention. Par exemple, l'évaluation du programme de transfert conditionnel de fonds Progres/Oportunidades au Mexique⁴ portait sur les taux de fréquentation scolaire plutôt que sur les résultats d'apprentissage.

Cependant, les évaluations d'impact réalisées trop tôt donnent une image inexacte des impacts. Dans certains cas, ils seront sous-estimés, n'ayant pas eu le temps de se concrétiser (par exemple, des enfants qui achèvent leur cursus scolaire après avoir participé à un programme d'intervention précoce). Dans d'autres cas, ils seront surévalués (par exemple, les compétences et connaissances de participants peuvent diminuer au fil du temps s'ils n'ont pas l'occasion de les exploiter et par conséquent de les maintenir à niveau). Dans de tels cas, il sera nécessaire de réaliser ultérieurement une étude de suivi pour examiner la durabilité de ces impacts.

3. DÉTERMINATION DES DISPOSITIFS EN MATIÈRE DE GOUVERNANCE ET DE GESTION

Une évaluation d'impact doit être planifiée et gérée de façon officielle, et constituer un projet à part entière faisant l'objet de dispositifs de gestion clairement décrits dès le début du processus.

Les structures et processus importants sont notamment :

- **équipe de gestion de l'évaluation ou comité de pilotage**, pour sélectionner et superviser le travail des évaluateurs externes engagés pour réaliser l'évaluation, notamment en examinant les TdR, le rapport initial et le projet de rapport final. L'équipe fournit des conseils techniques aux évaluateurs, gère les aspects opérationnels de l'évaluation, comme le budget, les visites de terrain et les échanges avec les bureaux de pays et sections de l'UNICEF concernés, et s'occupe de l'assurance qualité des outils d'évaluation (p. ex. la pertinence d'une enquête spécifique). Dans une évaluation conjointe, cette équipe comprend des membres du personnel de l'UNICEF et des représentants de chaque organisation partenaire ;

³ Pour de plus amples renseignements, consulter : Politique d'évaluation révisée de l'UNICEF, E/ICEF/2013/14, Session annuelle du Conseil d'administration, 18–21 juin 2013, UNICEF, 2013.

⁴ Voir Secretaría de Desarrollo Social (SEDESOL), Oportunidades, www.oportunidades.gob.mx (en espagnol).

- **groupe de référence de l'évaluation**, pour apporter des conseils et orientations techniques relatives à une interaction culturellement adaptée aux communautés locales (y compris des outils et méthodes d'évaluation adaptés à la culture locale), à l'élaboration des TdR, aux critères de sélection pour l'équipe d'évaluation, à l'examen du plan d'évaluation et du projet de rapport d'évaluation. Le groupe de référence peut être composé de spécialistes internes et externes (y compris des parties prenantes clés).

L'approbation du rapport final d'évaluation s'effectue généralement à l'issue d'un processus d'examen par l'équipe de gestion et le groupe de référence de l'évaluation, suivi de l'approbation/validation finale par le chef de bureau de l'UNICEF.

Les principaux destinataires de l'évaluation d'impact (organisations ou particuliers) doivent être identifiés et impliqués autant que possible dans le processus d'évaluation, soit au sein de l'équipe de gestion de l'évaluation, soit au sein d'un groupe de référence des parties prenantes. Il peut s'agir de responsables et membres du personnel de l'UNICEF au niveau local, régional et international ou de représentants des organisations partenaires locales, communautés locales, organisations non gouvernementales (ONG) locales, donateurs locaux, ministères partenaires, bureaux mondiaux des organisations et ONG partenaires, bureaux mondiaux des donateurs et instituts de recherche pertinents.

Il est nécessaire de mettre en place des processus appropriés afin de comprendre ce que les principaux destinataires attendent de l'évaluation et de déterminer quelles attentes sont prioritaires. D'autres destinataires (p. ex. des organismes mettant en œuvre des programmes similaires) peuvent en outre souhaiter en savoir plus sur l'évaluation d'impact, mais leurs besoins ne constituent toutefois pas nécessairement une priorité pour l'évaluation.

4. PRÉPARATION DE LA MISE EN ŒUVRE D'UNE ÉVALUATION D'IMPACT

Étude de la théorie du changement

Une théorie du changement explique comment les activités sont comprises afin de produire une série de résultats qui contribuent à la réalisation des effets escomptés ou observés. Il est probable qu'une théorie du changement ait été élaborée lors de la phase de planification d'un programme ou d'une politique. Lors de la préparation d'une évaluation d'impact, la théorie du changement doit être examinée et révisée au besoin. Par exemple, la théorie du changement en vigueur peut présenter des lacunes ou des hypothèses irréalistes qui devraient être révisées ; il se peut également qu'elle ne soit plus à jour si le programme a évolué depuis sa formulation.

Les versions d'une théorie du changement n'expliquent pas toutes *comment* une intervention est censée fonctionner (p. ex. chaînes de résultats et cadres logiques). Un programme peut par exemple comporter une théorie du changement qui précise que des activités d'éducation en matière d'hygiène seront menées afin d'expliquer l'importance du lavage des mains pour la santé, dans le but de promouvoir le lavage des mains et d'améliorer la santé. Une théorie du changement efficace explique le lien de causalité qui est censé être en place ; c'est-à-dire l'éducation en matière d'hygiène fonctionne-t-elle en apportant de nouvelles informations (p. ex. que le savon est nécessaire pour améliorer l'efficacité du lavage des mains, ou qu'il est possible d'utiliser des cendres en remplacement du savon, etc.), en rappelant à la population des informations qu'elle connaît déjà, en renforçant ou en modifiant les normes sociales, ou encore en créant des opportunités de discussion et de réduction des obstacles à l'accès à l'eau et au savon ? Lors de la préparation d'une évaluation d'impact, il peut s'avérer utile de préciser ce lien de causalité ou d'établir qu'il s'agit d'une importante

lacune en matière d'information. L'évaluation d'impact peut être de meilleure qualité si elle s'appuie sur la théorie du changement pour :

- identifier les variables pertinentes devant être incluses dans la collecte de données ;
- identifier les résultats intermédiaires pouvant être utilisés comme marqueurs de succès, dans des situations où les impacts d'intérêt surviennent après la période d'évaluation ;
- identifier les aspects de la mise en œuvre qui doivent être examinés afin de vérifier si la non réalisation des effets escomptés est due à l'impossibilité de mettre en œuvre efficacement l'intervention ;
- identifier les facteurs contextuels potentiellement pertinents devant être inclus dans la collecte de données ;
- guider l'analyse des données ;
- fournir un cadre de travail pour rendre compte des résultats.

De plus amples informations sur l'élaboration et l'utilisation d'une théorie du changement sont disponibles dans la Note N°2, Théorie du changement.

Identification et mobilisation des ressources

Il est important de prévoir dès le début une « enveloppe d'évaluation » (c'est-à-dire les fonds et les ressources nécessaires à l'évaluation) de sorte que des ressources suffisantes soient disponibles lors de sa réalisation. Si les décisions relatives à la conception et aux coûts de l'évaluation au sens large peuvent être prises pendant la phase de planification de l'intervention, des ressources suffisantes peuvent être réservées aux activités d'évaluation. Une estimation du montant nécessaire peut être obtenue en consultant les budgets de précédentes évaluations analogues, en utilisant des modèles ou en élaborant un modèle initial et en calculant les coûts correspondants⁵.

Outre les fonds finançant éventuellement l'équipe d'évaluation externe, il est également important de veiller à ce que les principaux contributeurs, notamment le groupe de gouvernance, les organismes d'examen (tels qu'un groupe consultatif technique ou des pairs examinateurs), les [informateurs clés](#) et les organisateurs de visites sur site, aient suffisamment de temps à consacrer à l'évaluation.

Il est également nécessaire d'avoir accès aux systèmes de données et à la documentation existante. Il est important de bien comprendre les données d'ores et déjà disponibles afin que l'évaluation soit conçue de manière à vérifier les informations et à combler les lacunes plutôt qu'à recueillir des données redondantes.

Détermination du processus d'élaboration de la méthode d'évaluation

Pour mettre en place une méthode efficace, il est nécessaire de bien comprendre les besoins et ressources spécifiques de l'évaluation, et de bénéficier d'une expertise technique générale. Il est par conséquent préférable de concevoir l'évaluation à l'aide d'un processus qui réunit tous ces éléments. Le schéma ci-dessous présente les quatre processus envisageables pour la mise au point de la méthode d'évaluation.

⁵ Pour obtenir une liste de contrôle des éléments à prendre en compte, consulter Horn, Jerry, « A Checklist for Developing and Evaluating Evaluation Budgets », Western Michigan University Checklist Project, Kalamazoo, 2001. Voir http://www.wmich.edu/evalctr/archive_checklists/evaluationbudgets.pdf (en anglais).

Figure 1. Processus envisageables pour l'élaboration de la méthode d'évaluation

Méthode élaborée par l'organisme commanditaire, en amont	Méthode initiale élaborée par l'organisme commanditaire, puis révisée	Méthode initiale élaborée par des évaluateurs, puis révisée	Méthode élaborée par des évaluateurs, en amont sous forme de projet distinct
<ul style="list-style-type: none"> Énoncée dans les TdR 	<ul style="list-style-type: none"> Méthode initiale révisée dans le rapport initial 	<ul style="list-style-type: none"> Méthode initiale révisée dans le rapport initial 	<ul style="list-style-type: none"> Énoncée dans les TdR L'évaluation en elle-même est commandée dans le cadre d'un projet distinct

Méthode élaborée par l'organisme commanditaire : c'est là que les TdR précisent le modèle, soit en le nommant de façon explicite, soit en lui réservant une ligne budgétaire (p. ex., X nombre de jours consacrés à l'examen des documents en interne et aux entretiens sur le terrain). Cette solution peut être appropriée si l'organisme commanditaire dispose d'une expertise technique suffisante, en raison des compétences de son personnel ou parce que l'évaluation s'appuie sur un modèle préexistant. Sinon, il se peut que cette solution ne permette pas d'utiliser au mieux les opportunités de débouchés en raison du nombre limité de modèles et méthodes d'évaluation connues par les personnes mandatées.

Méthode initiale élaborée par l'organisme commanditaire, puis révisée : c'est là que les TdR énoncent un modèle indicatif qui est ensuite révisé par les évaluateurs tout de suite après leur nomination. Dans un rapport initial, les évaluateurs présentent le modèle final, y compris les éventuelles révisions, basées sur les discussions avec l'organisme commanditaire et un examen des données disponibles. Cette option permet l'ajout de contributions techniques et logistiques supplémentaires à la méthodologie adoptée, qui est par conséquent généralement plus solide que le modèle indicatif proposé initialement.

Méthode initiale élaborée par des évaluateurs, puis révisée : c'est là que les TdR prévoient que les évaluateurs potentiels proposent et chiffrent un modèle d'évaluation. Ces éléments sont ensuite examinés puis révisés à la suite de leur nomination. Dans un rapport initial, les évaluateurs présentent le modèle final, y compris les éventuelles révisions, basées sur les discussions avec l'organisme commanditaire et un examen des données disponibles. Comme pour l'option précédente, ce scénario devrait permettre d'élaborer un modèle plus solide, car il s'appuie sur une meilleure compréhension de la situation sur le terrain.

Méthode élaborée sous forme de projet distinct : c'est là qu'un projet distinct est entrepris pour élaborer la méthode d'évaluation. Ce processus peut nécessiter une vaste consultation des destinataires de l'évaluation et un examen des données disponibles sous forme d'une étude d'évaluabilité⁶. La réalisation de l'évaluation elle-même fait alors l'objet d'un contrat distinct ; l'équipe à l'origine de la méthodologie ne sera pas nécessairement choisie pour réaliser l'évaluation si elle ne dispose pas de l'expertise, de la couverture ou de la disponibilité nécessaires. Cette option est la plus

⁶ Pour de plus amples informations, consulter la page Internet de BetterEvaluation « Evaluability Assessment », http://betterevaluation.org/themes/evaluability_assessment (en anglais).

appropriée lorsqu'une méthode complexe s'avère indispensable pour répondre aux caractéristiques particulières de l'évaluation.

5. ÉLABORATION DES TERMES DE RÉFÉRENCE (TDR) POUR L'ÉVALUATION

L'élaboration de TdR adaptés représente une étape cruciale du processus d'évaluation d'impact. Si le format des TdR peut varier, en particulier s'il s'agit d'une évaluation conjointe, ils doivent toutefois présenter les éléments essentiels de l'évaluation d'impact, notamment son but, ses objectifs et utilisations prévues, les questions clés de l'évaluation (*Key Evaluation Questions*, KEQ) de haut niveau auxquelles elle cherche à répondre, les livrables et échéanciers spécifiques concernés et les compétences essentielles attendues de l'équipe d'évaluation. Les TdR présentent généralement la méthode proposée pour l'évaluation d'impact – les options pour la développer et la perfectionner ont été abordées ci-dessus.

Contexte

Les TdR doivent fournir un résumé du déroulement de l'intervention à évaluer, préciser ce qui a suscité la décision d'entreprendre une évaluation d'impact et définir les concepts principaux : par exemple, les TdR pour une évaluation d'impact des systèmes de préparation aux situations d'urgence de l'UNICEF incluaient une définition précise de la notion de « préparation »⁷.

But, objectifs et portée de l'évaluation (dont les questions clés de l'évaluation)

Le but de l'évaluation répond à la question « pourquoi entreprendre une évaluation d'impact ? ». Les évaluations entreprises pour encourager l'acquisition de connaissances doivent préciser clairement le public visé par l'apprentissage en question, comment ce public sera impliqué dans le processus d'évaluation afin de garantir sa pertinence et sa crédibilité, et s'il existe des décisions à prendre quant aux domaines auxquels appliquer ces connaissances. Les évaluations qui sont effectuées pour appuyer la responsabilisation doivent définir clairement qui est responsable, vis-à-vis de qui et pourquoi.

Les objectifs de l'évaluation sont habituellement exprimés selon les [critères d'évaluation](#)⁸ de l'OCDE-CAD (voir tableau 2).

⁷ Bureau de l'évaluation de l'UNICEF, *Evaluation of UNICEF's Emergency Preparedness Systems, Terms of Reference – Final*, UNICEF, 2012. Voir http://www.unicef.org/evaluation/files/Evaluation_of_UNICEF_Emergency_Preparedness_Systems_-_FINAL_ToR.pdf (en anglais).

⁸ Organisation de coopération et de développement économiques – Comité d'aide au développement, *Évaluation des programmes de développement, DAC Criteria for Evaluating Development Assistance (Critères du CAD pour l'évaluation de l'aide au développement)*, page Internet, OCDE. Voir <http://www.oecd.org/fr/cad/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm> (en anglais).

Tableau 2. Critères d'évaluation de l'OCDE-CAD

Pertinence	Mesure dans laquelle les objectifs de l'intervention correspondent aux attentes des bénéficiaires, aux besoins du pays, aux priorités internationales et aux politiques des partenaires.
Efficacité	Mesure dans laquelle les objectifs de l'intervention ont été atteints, ou devraient être atteints, en tenant compte de leur importance relative.
Efficiences	Mesure de l'efficacité de l'usage des ressources/intrants (fonds, expertise, temps, équipements, etc.) pour l'obtention de résultats.
Impact	Effets à long terme, positifs et négatifs, primaires et secondaires, induits par l'intervention, directement ou non, intentionnellement ou non.
Durabilité	Maintien des bénéfices résultant de l'intervention après l'interruption de l'aide au développement. La durabilité concerne la probabilité d'obtenir des bénéfices sur le long terme.

Les évaluations d'impact peuvent aussi être destinées à étudier l'équité, l'égalité entre les sexes et les droits de l'homme.

Une évaluation peut être axée sur une partie seulement de ces critères. Une évaluation d'impact doit nécessairement décrire plus spécifiquement chacun des critères afin de déterminer clairement « à quoi l'on reconnaît la réussite ». De plus amples informations sur la définition des critères d'évaluation sont disponibles dans la Note N°3, Critères d'évaluation.

Les évaluations d'impact doivent s'attacher à répondre à un petit nombre de questions clés d'évaluation de haut niveau. Il ne s'agit pas de questions à poser lors d'un entretien, mais plutôt de questions auxquelles un ensemble de données probantes permettront de répondre (voir les exemples du tableau 3).

Ces questions doivent être en rapport direct avec les utilisations prévues de l'évaluation d'impact. Si une évaluation est par exemple destinée à éclairer une décision relative à l'extension d'un programme pilote, il est alors évident que les questions « Cela a-t-il fonctionné ? » ou « Quels ont été les impacts ? » ne sont pas suffisantes. Il est nécessaire de bien comprendre quelles activités et quels facteurs contextuels favorables ont permis d'obtenir ces impacts, afin de reproduire les effets positifs d'un programme pilote réussi. Les préoccupations relatives à l'équité exigent également que les évaluations d'impact de l'ensemble des programmes de l'UNICEF ne se cantonnent pas à déterminer simplement l'impact moyen, mais plutôt qu'elles identifient pour qui ces programmes ont effectivement fonctionné et comment.

Tableau 3. Exemples de questions clés posées lors d'évaluations d'impact relatives aux critères d'évaluation de l'OCDE-CAD

Pertinence

Dans quelle mesure les impacts prévus correspondent-ils aux besoins des participants visés ?

Dans quelle mesure les impacts prévus correspondent-ils aux priorités énoncées des gouvernements partenaires et du système des Nations Unies ?

Efficacité

Est-ce que l'intervention a produit [les impacts prévus] à court, moyen et long terme ?

Pour qui, de quelle manière et dans quelles circonstances [l'intervention] a-t-elle fonctionné ?

Les [impacts prévus] ont-ils atteint tous les participants visés ?

Comment [l'intervention] a-t-elle contribué aux [impacts prévus] ?

Quelles caractéristiques particulières de [l'intervention] ont eu un impact ?

Quels écarts par rapport aux prévisions se sont manifestés au cours de la mise en œuvre ?

Quelle a été la qualité de la mise en œuvre sur les différents sites ? Dans quelle mesure et de quelle façon la mise en œuvre a-t-elle changé au fil du déroulement de l'intervention ? Dans quelle mesure les différences d'impact peuvent-elles être expliquées par les écarts constatés dans la mise en œuvre ?

Comment l'intervention s'est-elle articulée avec d'autres pour aboutir aux résultats obtenus ?

Qu'est-ce qui a aidé l'intervention à produire ces impacts ? Qu'est-ce qui a empêché l'intervention de produire ces impacts ?

Efficiences

Quelles ressources ont été utilisées pour produire ces résultats ? En quoi diffèrent-elles de celles d'interventions similaires ?

Quelles stratégies ont été utilisées pour assurer l'efficacité de l'intervention ?

Impact

Quels impacts imprévus (positifs et négatifs) l'intervention a-t-elle produits ?

Durabilité

Les impacts sont-ils susceptibles de s'inscrire dans la durée ? (à étudier à la fin de l'intervention)

Les impacts ont-ils été durables ? (étude de suivi requise)

Les TdR doivent énoncer clairement le champ de l'évaluation. Par exemple, les TdR pour l'évaluation du programme conjoint UNFPA-UNICEF sur les mutilations génitales féminines/excision : Accélérer le changement⁹ faisaient explicitement référence aux différents niveaux du programme et précisaient dans quelle mesure l'évaluation devait étudier chacun d'entre eux.

⁹ Fonds des Nations Unies pour la population et Fonds des Nations Unies pour l'enfance, Évaluation conjointe du programme conjoint UNFPA-UNICEF sur les mutilations génitales féminines/excision : Accélérer le changement, 2008–2012, Volume I, UNFPA/UNICEF, New York, septembre 2013. Voir

Conception et méthodes de l'évaluation

La méthode d'évaluation définit comment répondre aux questions clés de l'évaluation. Elle spécifie les modèles pour l'attribution causale, y compris l'éventuelle constitution de groupes témoins et leur mode de constitution, ainsi que les méthodes de collecte et d'analyse des données. Il est recommandé d'élaborer un tableau détaillant les méthodes de collecte/regroupement et d'analyse de données (y compris en utilisant des données existantes) qui seront utilisées pour répondre à chaque question clé de l'évaluation.

Il est utile d'identifier, parmi les questions clés de l'évaluation, les différents types de questions – descriptive, causale et évaluative. Les réponses des évaluations d'impact à toutes ces questions doivent être crédibles : en quoi les choses ont-elles changé, dans quelle mesure ce changement peut-il être imputé à l'intervention évaluée et, dans l'ensemble, l'intervention est-elle réussie ? La plupart des évaluations d'impact abordent à la fois les questions descriptives et les questions causales, mais bon nombre d'entre elles ne traitent pas les questions évaluatives (relatives à la qualité et à la valeur). C'est pourtant précisément cela qui justifie l'utilité et la pertinence des évaluations par rapport à une simple mesure [d'indicateurs](#) ou de simples résumés d'observations et d'histoires (voir Note N°4, Logique d'évaluation).

Il est nécessaire d'avoir recours à différents modèles et méthodes de collecte et d'analyse des données d'évaluation pour répondre à ces divers types de questions d'évaluation, énumérés ci-après.

Questions descriptives

Les questions descriptives portent sur la description de la situation, notamment de la situation initiale et de son évolution, des activités de l'intervention et d'autres programmes ou politiques connexes, du contexte relatif aux caractéristiques des participants et de l'environnement de mise en œuvre.

Diverses méthodes de collecte/recueil de données permettent de répondre, entre autres, aux questions descriptives :

- la collecte de données auprès d'individus ou de groupes, par le biais d'entretiens, de questionnaires ou méthodes spécialisées telles que la [dotmocracy](#) (méthode participative d'aide à la décision), le [classement hiérarchique par cartes](#), les calendriers saisonniers, les techniques projectives et les histoires ;
- l'observation structurée ou non structurée, avec ou sans participants, retranscrite à l'aide de notes, de photos ou de vidéos, participative ou non participative ;
- les mesures physiques : mesures biophysiques ou informations géographiques ;
- les documents et données existants, tels qu'ensembles de données, statistiques officielles, dossiers concernant les projets et données des médias sociaux disponibles.

La plupart du temps, l'association efficace de données quantitatives et qualitatives permet de fournir une image plus précise de la situation. De nombreuses évaluations d'impact ne rassemblent pas les [données de référence](#) appropriées relatives aux conditions de l'intervention ou les données sur sa mise en œuvre.

De plus amples informations sur les différentes options pour la collecte et l'analyse des données et sur la mise en commun de données qualitatives et quantitatives sont disponibles dans la Note N°10, Présentation des méthodes de collecte et d'analyse de données dans l'évaluation d'impact. Des conseils spécifiques sur le recours à des entretiens sont disponibles dans la Note N° 12, Entretiens.

Questions causales

Les questions causales examinent dans quelle mesure les changements observés résultent de l'intervention évaluée ou d'autres facteurs (autres programmes et/ou politiques).

Il existe trois modèles possibles traitant de l'attribution causale :

- les modèles expérimentaux, qui constituent un groupe contrôle par randomisation ;
- les modèles quasi expérimentaux, qui constituent un groupe témoin par correspondance, régression avec discontinuité, scores de propension ou autre méthode ;
- les modèles non expérimentaux, qui examinent systématiquement si les données probantes correspondent ou non aux attentes dans le cas où l'intervention produit effectivement des impacts, et si d'autres facteurs peuvent également l'expliquer.

Toutes ces options nécessitent un investissement important dans la préparation et la collecte précoce de données, et ne peuvent être employées si l'évaluation d'impact se limite à un bref exercice mené vers la fin de la mise en œuvre de l'intervention – un point faible fréquemment rencontré dans bon nombre d'évaluations d'impact.

De plus amples informations sur les différentes méthodes possibles sont disponibles dans la Note N°6, Présentation des stratégies d'attribution causale ; Note N°7, Essais contrôlés randomisés ; Note N°8, Modèles et méthodes quasi expérimentaux et Note N°9, Études de cas comparatives.

Questions évaluatives

Les questions évaluatives portent sur la conclusion générale, à savoir si l'on peut considérer le programme ou la politique comme une réussite, une amélioration ou la meilleure option.

Étant donné que les interventions de développement ont souvent de nombreux impacts inégalement répartis (l'équité est une préoccupation particulière du système des Nations Unies), ces questions représentent un élément essentiel de l'évaluation d'impact. Par exemple, un programme de développement économique doit-il être considéré comme une réussite s'il génère une augmentation du revenu des ménages, mais produit en parallèle des impacts dangereux sur l'environnement ? Doit-il être considéré comme une réussite si le revenu moyen des ménages augmente mais si celui des ménages les plus pauvres diminue ?

Dans toute évaluation d'impact, il est primordial de commencer par définir ce que l'on entend par « réussite » (qualité, intérêt). Il est pour cela possible d'utiliser une [grille critériée](#) spécifique qui définit les différents niveaux de performance (ou normes) pour chaque critère d'évaluation, en déterminant les données probantes devant être recueillies et synthétisées pour parvenir à des conclusions justifiables quant à l'intérêt de l'intervention. Il faut au minimum que soit clairement précisé quelles concessions sont adaptées pour pondérer les divers impacts ou effets distributionnels. C'est un point faible : de nombreuses évaluations d'impact ne traitent pas cette question de façon systématique, voire ne l'abordent pas du tout. De plus amples informations sur l'importance d'aborder les questions évaluatives et sur les manières spécifiques de les adresser (processus recommandé et moyens de parvenir à des conclusions justifiables) sont disponibles dans la Note N°4, Logique d'évaluation.

Profil de l'équipe d'évaluation

Les TdR doivent déterminer les critères de sélection de l'équipe d'évaluation, y compris les compétences obligatoires et souhaitées, par exemple : expérience professionnelle dans le pays ou la région en question ; compétences linguistiques ; connaissances fondamentales (dans un secteur ou un thème particulier, comme les questions sexospécifiques) ; compétences techniques en conception d'évaluation, collecte de données, analyse de données et rédaction de rapports ; compétences en

communication et en gestion de projet. Il peut s'avérer nécessaire de veiller à ce que l'ensemble de l'équipe regroupe les compétences adéquates, par exemple connaissances à jour des réseaux locaux, capacité d'accès aux informateurs clés et connaissance d'évaluations d'interventions comparables dans d'autres régions. Il est important de s'assurer que l'équipe dispose des compétences techniques nécessaires pour la collecte et l'analyse des données, des compétences en communication nécessaires pour négocier efficacement la conception et rendre compte des résultats, et qu'elle comprenne que les évaluations doivent être axées sur la production de résultats utiles et accessibles, et non de rapports académiques.

Si des approches participatives impliquant des membres de groupes de parties prenantes, notamment les membres de la communauté, sont jugées appropriées pour réaliser les évaluations, l'équipe d'évaluation doit alors posséder l'expérience et les compétences nécessaires pour faciliter cette participation.

De plus amples informations sur les stratégies de mobilisation de différentes parties prenantes (y compris les membres de la communauté) sont disponibles dans la Note N°5, *Approches participatives*.

D'autres informations sur les processus permettant d'élaborer des TdR en vue d'une évaluation sont disponibles dans la Note technique N°2 sur l'évaluation de l'UNICEF, *What goes into a Terms of Reference (TOR)*¹⁰ (Que constitue un Terme de référence).

6. CONSTITUTION DE L'ÉQUIPE D'ÉVALUATION

La constitution d'une équipe d'évaluation externe doit suivre les procédures standard de l'UNICEF conçues pour garantir la disponibilité d'un large panel d'évaluateurs potentiels à partir duquel choisir la ou les personnes les plus appropriées.

Les TdR doivent fournir des informations suffisantes pour permettre aux évaluateurs potentiels de soumettre une proposition technique et financière étayée. C'est la raison pour laquelle il est préférable de mener les études exploratoires ou d'évaluabilité avant de décider de réaliser une évaluation d'impact et de rechercher des évaluateurs externes. Pour motiver un nombre important d'évaluateurs potentiels, il convient de publier les TdR à grande échelle, y compris par le biais de listes de discussion sur l'évaluation et de courriers adressés directement aux évaluateurs potentiels. Faire savoir qu'il existe des opportunités de rédaction de publications académiques peut encourager les évaluateurs et chercheurs universitaires à soumettre des propositions. Toute stratégie efficace pour sélectionner la meilleure équipe d'évaluation commence par la définition de critères de sélection adaptés dans les TdR. Étudier des exemples de rapports d'évaluation, dans la mesure du possible, et discuter avec les équipes potentielles peut aider à mieux cerner dans quelle mesure leurs compétences et valeurs correspondent aux exigences de l'évaluation.

Il est important de laisser suffisamment de temps entre la diffusion des TdR et la date limite de soumission des propositions, et entre l'acceptation d'une proposition et le début de la mission. Les délais pour l'achèvement des travaux doivent également être raisonnables. En effet, des délais trop serrés peuvent considérablement limiter le nombre d'évaluateurs en mesure de travailler sur le projet et compromettre la qualité du travail.

¹⁰ Bureau de l'évaluation de l'UNICEF, « What goes into a Terms of Reference (TOR) », *Note technique n° 2 sur l'évaluation*, UNICEF, 2002. Il s'agit d'une liste des questions que doivent prendre en compte les responsables de l'évaluation de l'UNICEF lors de l'élaboration des termes de référence. Voir http://www.unicef.org/evaluation/files/TechNote2_TOR.pdf (en anglais).

7. SUPERVISION DE L'ÉVALUATION

Plan de travail d'évaluation

Dans bien des cas, les TdR présentent un plan de travail indicatif, qui précise les différentes étapes de la collecte et de l'analyse des données ainsi que les dates limites pour les différents livrables. Dans la mesure du possible, le plan de travail doit prévoir la rédaction de rapports intermédiaires et d'analyses régulières, plutôt que de tous les réserver pour la fin de l'évaluation. Ce plan de travail doit être examiné et révisé dans le cadre du [rapport initial](#).

Une fois accepté, la direction doit surveiller que le plan de travail et les structures de gouvernance sont effectivement en place – et révisés, le cas échéant – afin de tenir compte des problèmes émergents ou retards imprévus.

Il convient de prévoir le temps nécessaire à l'examen et la révision des livrables clés, y compris la méthode et le plan de travail révisés, les outils d'analyse et de collecte de données (s'ils doivent être approuvés), ainsi que les rapports préliminaires.

Communication des résultats

Les rapports devraient être rédigés conformément au GEROS¹¹ (système mondial de contrôle des évaluations) développé par l'UNICEF, et les évaluateurs doivent avoir à leur disposition un modèle pour faciliter la publication d'un document conforme.

Les rapports doivent satisfaire aux règles UNICEF-Adapted UNEG Evaluation Reports Standards¹² (règles relatives aux rapports d'évaluation du GNUE adaptées par l'UNICEF). Ces règles visent globalement à garantir que :

1. le rapport est bien structuré, logique, clair et complet ;
2. le rapport offre une description claire et exhaustive de l'« objet » de l'évaluation ;
3. le but, les objectifs et la portée de l'évaluation sont parfaitement expliqués ;
4. le rapport donne une description transparente de la méthode appliquée pour l'évaluation, qui explique clairement comment l'évaluation a été spécifiquement conçue pour répondre aux critères d'évaluation, générer des réponses aux questions d'évaluation et atteindre les buts de l'évaluation ;
5. les résultats observés répondent directement aux critères d'évaluation et questions détaillées dans la section du rapport relative à la portée et aux objectifs, et s'appuient sur des données probantes issues des méthodes de collecte et d'analyse des données décrites dans la section du rapport relative à la méthode ;
6. les conclusions présentent des jugements raisonnables fondés sur les résultats observés et étayés par des données probantes, et apportent des informations concernant l'objet et le but de l'évaluation ;
7. les recommandations sont pertinentes quant à l'objet et au but de l'évaluation, s'appuient sur des données probantes et conclusions, et ont été élaborées avec la participation des parties prenantes compétentes ;

¹¹ Pour de plus amples informations sur les exigences relatives aux rapports, consulter : Fonds des Nations Unies pour l'enfance, Global Evaluation Reports Oversight System (GEROS), UNICEF, 2013. Voir http://www.unicef.org/evaluation/files/GEROS_Methodology_v7.pdf (en anglais).

¹² Bureau de l'évaluation de l'UNICEF, UNICEF-Adapted UNEG Evaluation Reports Standards, UNICEF, 2010. Voir http://www.unicef.org/evaldatabase/files/UNEG_UNICEF_Eval_Report_Standards.pdf (en anglais).

8. le rapport illustre dans quelle mesure la conception et la mise en œuvre de l'objet, l'évaluation des résultats et le processus d'évaluation intègrent les aspects relatifs à l'égalité des sexes et l'approche axée sur les droits de l'homme, y compris sur les droits de l'enfant.

Le rapport d'évaluation doit être structuré autour des messages clés du résumé exécutif et permettre d'accéder facilement à des explications et à des données probantes plus détaillées. L'utilisation d'observations synthétiques comme intitulés, et de titres pour les graphiques et tableaux permet aux lecteurs d'être plus attentifs à ces messages clés.

Réponse et mesures correctives de la direction

Une évaluation doit faire l'objet de mesures correctives formelles de la direction¹³ sous la forme d'une réponse écrite aux recommandations du rapport d'évaluation. Certains organismes intègrent ces mesures dans le rapport final ; à l'UNICEF, les mesures correctives proposées par la direction se présentent sous la forme d'un document distinct. Ce retour peut indiquer l'accord, l'accord partiel ou le désaccord de la direction vis-à-vis d'une recommandation et doit apporter une explication en cas d'acceptation partielle ou de rejet. Pour les recommandations qui ont été acceptées, ou partiellement acceptées, il convient de déterminer des actions clés de suivi, d'en préciser le délai de réalisation et d'indiquer le nom de l'unité responsable.

Il est important de désigner une personne qui sera chargée de coordonner la réponse de la direction et de convenir d'un délai pour la réception des commentaires. La direction dispose de deux mois à compter de la réception du rapport d'évaluation final¹⁴ pour y répondre.

8. DONNER SUITE À L'ÉVALUATION

Diffusion des résultats

Outre le rapport d'évaluation formel, il est possible d'avoir recours à d'autres processus et produits de connaissances afin de rendre les résultats plus facilement accessibles. Il s'agit notamment d'autres documents écrits, comme une note sur l'évaluation, un e-bulletin ou une entrée sur un wiki, comportant tous un résumé de l'évaluation.

Les constatations peuvent aussi être partagées à travers d'activités, telles que des présentations lors de réunions, d'ateliers et dans des communautés de pratique existantes et pour des séminaires. Ces présentations peuvent même être enregistrées et partagées plus largement, par exemple, sur YouTube ou Vimeo.

Toutes les évaluations de l'UNICEF doivent également être téléchargées dans la base de données sur les évaluations de l'UNICEF dans le cadre du système de rapports annuels des bureaux de pays (RABP).

¹³ Fonds des Nations Unies pour l'enfance, *Guidance for Management Response to Evaluation: Enhancing Critical Engagement for the Strategic Use of Evaluations*, UNICEF, 2012. Voir http://www.unicef.org/evaldatabase/files/2011_South_Africa_-_management_response_SCCFS_evaluation.pdf (en anglais), et <https://intranet.unicef.org/epp/evalsite.nsf/8e1ddc662803020785256ede00706595/6a1a69ce471c376585257c40006677c1?OpenDocument> (en anglais).

¹⁴ Groupe des Nations Unies pour l'évaluation, *UNEG Guidance on Preparing Management Responses to UNDAF Evaluations*, document d'orientation, GNUE, 2014. Voir http://www.unevaluation.org/UNDAF_management_response_guidance (en anglais).

Supervision du suivi

Le progrès de la mise en œuvre des recommandations doit être supervisé dans le cadre des processus annuels de planification et d'examen. Il convient de rendre compte du niveau d'avancement de chacune des actions précédemment identifiées : « débuté », « non débuté », « terminé » ou « ne s'applique plus ».

Recenser les enseignements tirés au sujet de l'évaluation

Idéalement, les enseignements tirés de l'évaluation doivent également être documentés et partagés, au moins en interne. Plus précisément, la politique d'évaluation révisée de l'UNICEF¹⁵ stipule :

« *Les chefs de division* sont responsables de la planification, de la dotation en ressources et de l'attribution des évaluations des politiques et initiatives mondiales dont ils sont responsables. Il leur appartient aussi de tirer parti des enseignements et des recommandations pertinents des évaluations. Les chefs de division mettent en œuvre et soutiennent les actions visant à renforcer la capacité d'évaluation et l'implication des unités du siège qu'ils supervisent et des réseaux mondiaux qu'ils gèrent. »

Cette exigence est directement liée aux mesures correctives formelles de la direction mentionnées ci-dessus, notamment des mesures spécifiques devant être prises sur la base des constatations de l'évaluation.

9. ANTICIPER ET TRAITER LES QUESTIONS D'ÉTHIQUE

Les évaluations doivent suivre les *Ethical Guidelines for Evaluation*¹⁶ (directives éthiques pour l'évaluation) définies par le Groupe des Nations Unies pour l'évaluation (GNUE) et auxquelles adhère l'UNICEF. Ces directives établissent un code de conduite pour les évaluateurs ainsi que des obligations d'ordre éthique pour les personnes participant à l'évaluation.

Ces obligations comprennent :

- le respect de la dignité et de la diversité ;
- les droits ;
- la confidentialité ;
- la prévention des préjudices.

Ces obligations doivent être abordées 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 autres conventions relatives aux droits de l'homme.

¹⁵ Fonds des Nations Unies pour l'enfance, *Politique d'évaluation révisée de l'UNICEF*, paragraphe 37, UNICEF, New York, 2013. Voir http://www.unicef.org/french/about/execboard/files/2013-14-Revised_evaluation_policy-ODS-French.pdf.

¹⁶ Groupe des Nations Unies pour l'évaluation, *Ethical Guidelines for Evaluation*, GNUE, New York, 2007. Voir https://www.iom.int/jahia/webdav/site/myjahiasite/shared/shared/mainsite/about_iom/eva_techref/UNEG_ethical_guidelines.pdf (en anglais).

L'UNICEF exige que toutes les évaluations auxquelles elle participe respectent les lois du pays dans lequel se situent ses bureaux concernés et fassent ainsi l'objet d'un examen éthique par un « organisme impartial et qualifié »¹⁷, si cela est jugé nécessaire.

Il convient d'accorder une attention particulière aux questions éthiques pour les évaluations dans lesquelles certains informateurs sont des enfants, que ce soit en tant que collecteurs de données ou autres. En octobre 2013, le Bureau de recherche de l'UNICEF, Innocenti, a publié *Ethical Research Involving Children*, une ressource en ligne¹⁸ qui réunit des experts se penchant sur certaines questions éthiques clés concernant les enfants et la manière de traiter ces questions dans différents contextes de recherche (et d'évaluation).

De plus amples informations sur les questions éthiques sont disponibles dans la Note N°5, Approches participatives, et dans la Note N°10, Présentation des méthodes de collecte et d'analyse de données dans l'évaluation d'impact.

10. GARANTIR UNE ÉVALUATION D'IMPACT DE QUALITÉ

Pour assurer la qualité d'une évaluation d'impact, il faut trouver le juste équilibre entre plusieurs critères contradictoires : la validité¹⁹, l'utilité (en particulier le caractère opportun et la pertinence), la faisabilité et le respect de l'éthique. Un responsable peut par exemple avoir à choisir entre un modèle d'évaluation pouvant aboutir à des résultats très précis et exhaustifs et une autre permettant d'obtenir les informations adéquates à temps pour éclairer une décision particulière. Certains modèles mettent l'accent sur l'amélioration de la validité interne (la capacité à établir correctement une attribution causale) au détriment de la validité externe (l'aptitude à appliquer les constatations à d'autres contextes).

Les stratégies pour garantir une évaluation d'impact de qualité sont notamment :

- avoir connaissance des concessions faites et aborder de façon explicite les options disponibles ;
- impliquer des personnes diverses dans la gouvernance de l'évaluation pour veiller à ce que son axe principal et les décisions prises concernant les concessions soient appropriés ;
- faire participer un groupe consultatif technique ou des pairs examinateurs afin d'étudier le modèle de l'évaluation et les rapports préliminaires.

11. AUTRES NOTES DE CETTE SÉRIE

Cette note est la première d'une série de notes méthodologiques sur l'évaluation d'impact produites par l'UNICEF. Les autres notes de cette série sont les suivantes :

¹⁷ Fonds des Nations Unies pour l'enfance, *Politique d'évaluation révisée de l'UNICEF*, paragraphe 54(e), UNICEF, New York, 2013. Voir http://www.unicef.org/french/about/execboard/files/2013-14-Revised_evaluation_policy-ODS-French.pdf.

¹⁸ Les nouvelles ressources comprennent : l'International Charter for Ethical Research Involving Children (Charte internationale pour la recherche éthique impliquant des enfants) ; un site Internet, www.childethics.com (en anglais), spécifiquement conçu pour fournir un référentiel riche d'informations basées sur des données probantes, des ressources et des liens vers des articles de journaux afin d'orienter et d'améliorer la recherche impliquant les enfants et de fournir une plateforme pour une plus ample réflexion et un dialogue critiques ; et un recueil de questions éthiques et de défis, dont un recueil 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](#) »).

¹⁹ Différents types de validité sont importants : la validité conceptuelle fait référence à la capacité des mesures et d'autres outils de représenter fidèlement la réalité ; la validité interne se réfère à la capacité à établir correctement une attribution causale, c'est-à-dire à déterminer si les changements observés sont dus à l'intervention ; et la validité externe concerne la capacité à généraliser les résultats à d'autres populations ou sites.

Les fondamentaux

1. Présentation de l'évaluation d'impact
2. Théorie du changement
3. Critères d'évaluation
4. Logique d'évaluation
5. Approches participatives

Modèle de recherche

6. Présentation des stratégies d'attribution causale
7. Essais contrôlés randomisés
8. Modèles et méthodes quasi expérimentaux
9. Études de cas comparatives

Collecte et analyse des données

10. Présentation des méthodes de collecte et d'analyse de données dans l'évaluation d'impact
11. Élaboration et sélection de mesures du bien-être des enfants
12. Entretiens
13. Modélisation

12. PRINCIPALES LECTURES ET LIENS UTILES

Les autres notes de cette série fournissent également des liens et des titres de publications relatifs à certains aspects spécifiques de l'évaluation d'impact. Certaines introductions générales de l'évaluation d'impact sont proposées ci-dessous :

IE4D Group, « Impact Evaluation for Development: Principles for Action », IE4D Group, 2011. Cette publication aborde les stratégies à adopter pour gérer et réaliser l'évaluation du développement de façon à soutenir le développement. Voir <http://www.scalingimpact.net/files/Impact%20Evaluation%20for%20Development%20-%20Principles%20for%20Action.pdf> (en anglais).

Rogers, Patricia J., RMIT University et BetterEvaluation, « Introduction à l'évaluation d'impact », Notes sur l'évaluation d'impact, N°1 InterAction, Washington, 2012. Cette note fournit une vue d'ensemble des questions que doivent aborder les responsables de programme. Voir <http://www.interaction.org/sites/default/files/2%20-%20Introduction%20to%20Impact%20Evaluation%20-%20French.pdf>.

White, Howard, « Some Reflections on Current Debates in Impact Evaluation », *International Initiative for Impact Evaluation Working Paper*, N°1, 2009. Un résumé des principaux débats sur la mesure et l'attribution des impacts. Voir http://www.3ieimpact.org/media/filer/2012/05/07/Working_Paper_1.pdf (en anglais).

Bureau de l'évaluation de l'UNICEF, « What goes into a Terms of Reference (TOR) », Note technique N°2 sur l'évaluation, UNICEF, 2002. Il s'agit d'une liste des questions que doivent prendre en compte

les responsables de l'évaluation de l'UNICEF lors de l'élaboration des termes de référence. Voir http://www.unicef.org/evaluation/files/TechNote2_TOR.pdf (en anglais).

Bureau de l'évaluation de l'UNICEF, « Guidance on prioritization of major evaluations at the decentralized level », UNICEF, 2010, [https://intranet.unicef.org/epp/evalsite.nsf/0/089a45b4f5292588852577e00056f69c/\\$file/guidanceprioritizationofevaluations.pdf](https://intranet.unicef.org/epp/evalsite.nsf/0/089a45b4f5292588852577e00056f69c/$file/guidanceprioritizationofevaluations.pdf) (en anglais).

Bureau 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 University of Otago, Nouvelle-Zélande ; site Internet « Ethical Research Involving Children », 2013, www.childethics.com.

GLOSSAIRE

<u>Attribution causale</u>	<i>Confirmation d'une relation causale entre les changements observés (ou que l'on s'attend à observer) et une intervention spécifique.</i>
<u>Classement hiérarchique par cartes</u>	<i>Méthode de collecte de données qui permet de comprendre la manière dont les gens classent et hiérarchisent différents phénomènes</i>
<u>Critères d'évaluation</u>	<i>Valeurs (principes, attributs ou qualités considérées comme intrinsèquement bons, souhaitables, importants et présentant un intérêt général) qui seront utilisées dans une évaluation pour apprécier l'intérêt d'une intervention. Les critères de l'OCDE-CAD20, l'approche de la programmation fondée sur les droits de l'homme et les critères utilisés pour l'aide humanitaire constituent de bons exemples de critères d'évaluation.</i>
<u>Données de référence</u>	<i>Données initialement collectées pour servir de base de comparaison avec les données recueillies ultérieurement. Par exemple, les données collectées avant une intervention aux fins de comparaison avec les données recueillies après la mise en œuvre de l'intervention.</i>
<u>Dotmocracy (méthode participative d'aide à la décision)</u>	<i>Méthode de collecte de données auprès d'individus ou de groupes. Elle permet de recueillir les niveaux d'accord sur des déclarations écrites auprès d'un grand nombre de personnes.</i>
<u>Étude d'évaluabilité</u>	<i>Étude de faisabilité visant à déterminer s'il est possible d'évaluer un programme ou une politique. Elle permet de vérifier si une intervention est correctement définie, si ses résultats sont vérifiables et si l'évaluation est le meilleur moyen de répondre aux questions posées par les principales parties prenantes. Elle contrôle également la disponibilité des ressources adéquates pour réaliser une évaluation : données, budget, capacité et intérêt de l'utilisation des résultats.</i>
<u>Évaluation de processus</u>	<i>Évaluation de l'élaboration et de la mise en œuvre d'un programme/d'une intervention spécifique. Elle analyse notamment l'efficacité des opérations, la mise en œuvre, la prestation de services, etc. au sein d'un programme. L'évaluation continue du processus est appelée « suivi du programme ». Voir aussi : suivi.</i>
<u>Grille critériée</u>	<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>
<u>Groupe contrôle</u>	<i>Participants à une étude de recherche/évaluation ne bénéficiant pas du traitement/de l'intervention expérimental(e).</i>

²⁰ Comité d'aide au développement de l'Organisation de coopération et de développement économiques, Évaluation des programmes de développement, DAC Criteria for Evaluating Development Assistance (Critères du CAD pour l'évaluation de l'aide au développement), page Internet, OCDE. Voir <http://www.oecd.org/fr/cad/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>.

<u>Groupe témoin</u>	<i>Dans un modèle de recherche quasi-expérimental, il s'agit du groupe de participants à la recherche qui, à des fins de comparaison, ne bénéficie pas du traitement ou de l'intervention dont dispose le groupe expérimental. Généralement, les sujets du groupe témoin ne sont pas affectés de façon aléatoire à ce groupe, tout comme les sujets d'un groupe contrôle dans une méthode de recherche expérimentale. Voir aussi : groupe contrôle, groupe expérimental.</i>
<u>Indicateur</u>	<i>Mesure vérifiable choisie par l'équipe responsable d'un programme ou d'une politique en vue de prendre des décisions concernant le programme/la politique. Par exemple la proportion d'élèves ayant réussi un examen normalisé</i>
<u>Informateur clé</u>	<i>Participants à une recherche/évaluation susceptibles d'être bien informés sur un sujet et prêts à répondre sans parti pris.</i>
<u>Rapport initial</u>	<i>Rapport publié au terme de la période initiale qui souligne par exemple le travail déjà réalisé, les résultats préliminaires ainsi que les références et établit un plan de travail pour le reste du projet.</i>
<u>Suivi</u>	<i>Collecte et analyse des données au fur et à mesure de l'avancement du projet pour garantir l'adéquation de l'évaluation, de sa conception et des mesures de protection des participants.</i>
<u>Termes de référence (TdR)</u>	<i>Déclaration spécifiant le contexte, les objectifs, les destinataires, les questions clés d'évaluation, la méthodologie, les rôles et responsabilités, le calendrier, les livrables, les normes de qualité, les qualifications de l'équipe d'évaluation et d'autres aspects pertinents qui précisent les fondements d'un contrat passé entre l'UNICEF et les évaluateurs.</i>