

# L'ÉVALUATION EN CONTEXTE DE DÉVELOPPEMENT

An abstract graphic of a hand holding a glowing sphere. The hand is rendered as a white wireframe mesh, and the sphere is a glowing blue and green orb. The background is a dark blue gradient.

Enjeux, approches  
et pratiques

Sous la direction de  
Lynda Rey  
Jean Serge Quesnel  
Vénétia Sauvain

**JFD**  
Éditions

**ÉNAP** ÉCOLE NATIONALE  
D'ADMINISTRATION  
PUBLIQUE

# L'ÉVALUATION EN CONTEXTE DE DÉVELOPPEMENT

## Enjeux, approches et pratiques

Sous la direction de Lynda Rey, Jean Serge Quesnel et Vénétia Sauvain

Ce manuel est destiné aux personnes souhaitant s'initier à l'évaluation de programmes, en particulier en contexte de développement et de coopération internationale. À cet égard, tout en déroulant le fil d'une démarche évaluative classique, il présente plusieurs exemples en lien avec la réalité issue de la pratique des auteurs et autrices. Le public découvrira non seulement une synthèse des concepts et théories qui encadrent le domaine, mais sera également encouragé à découvrir des enjeux propres à l'environnement international.



**Lynda REY**, Ph. D., professeure en évaluation de programmes à l'École nationale d'administration publique et professeure associée au Centre Africain d'Études Supérieures en Gestion du Sénégal.



**Jean Serge QUESNEL**, ÉQ/CE, membre titulaire de la Société canadienne d'évaluation, professeur associé à l'École nationale d'administration publique et professeur invité à l'École du développement international et de la mondialisation de l'Université d'Ottawa.



**Vénétia SAUVAIN**, M. A., conseillère aux affaires internationales à l'École nationale d'administration publique (ENAP).

**Avec la collaboration de** Aref Ben Abdallah, Marie-Hélène Adrien, Eva Anstett, Oumoul Khayri Ba Tall, Ahmed Bencheikh, Faya François Bourouno, Isabelle Brunet, Diana Cartier, Guy Cauquil, Christian Dagenais, Thomas Delahais, Maxime Desfossés, Kris Essosola Eale, Thiery Tsou Fematouo, Alexandre Fortin, Christophe Franzetti, Mahmoud Ghouil, Anna Gueye, N'Diamé Gueye, Carla-Anide Guillaume, Deo-Gracias Houndolo, Tigawendé Prosper Kabore, Bobo Kabungu, Kadidiatou Kadio, Ousséni Kinda, Papa Souleymane Koulibaly, Julie Lane, Sylvie Lefrançois, Juan Manuel Lopez Mejia, Richard Marceau, Esther Mc Sween-Cadieux, Nicolas Montibert, Ada Ocampo, Miché Ouédraogo, Valéry Ridde, Élodie C. Roy, Abdoul Aziz Sall, Nebnoma Kader Arnaud Sawadogo, Mary Pat Selvaggio, Francine Sylvain, Michele Tarsilla, Serge Eric Yakeu Djiam et Wend-Yam Saturnin Zoetyande.

# L'évaluation en contexte de développement

Enjeux, approches et pratiques

Sous la direction de  
Lynda Rey, Jean Serge Quesnel et Vénétia Sauvain

**JFD**  
Éditions

L'évaluation en contexte de développement : enjeux, approches et pratiques  
Dirigé par Lynda Rey, Jean Serge Quesnel et Vénétia Sauvain

---

Catalogage avant publication de Bibliothèque et Archives nationales du Québec et Bibliothèque et Archives Canada

Titre : L'évaluation en contexte de développement : enjeux, approches et pratiques / dirigé par  
Lynda Rey, Jean Serge Quesnel et Vénétia Sauvain.

Identifiants : Canadiana 20220001421 | ISBN 9782897993016

Vedettes-matière : RVM : Administration publique – Évaluation. | RVM : Politique publique –  
Évaluation. | RVM : Évaluation de programme.

Classification : LCC JF1525.O74 E93 2022 | CDD 352.3/5—dc23

---

Les Éditions JFD inc.  
CP 15 Succ. Rosemont  
Montréal (Québec)  
H1X 3B6  
Courriel : [info@editionsjfd.com](mailto:info@editionsjfd.com)  
Web : [editionsjfd.com](http://editionsjfd.com)

Cet ouvrage est enregistré sous la licence *Creative Commons Attribution – Pas d'Utilisation  
Commerciale – Pas de Modification 4.0 International (CC BY-NC-ND 4.0)*  
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr>

ISBN : 978-2-89799-301-6

Dépôt légal : 2<sup>e</sup> trimestre 2022  
Bibliothèque et Archives nationales du Québec  
Bibliothèque et Archives Canada

Imprimé au Québec



# Table des matières

Remerciement.....	7
Mot du directeur général de l'École nationale d'administration publique .....	9
Mot du directeur des affaires internationales de l'École nationale d'administration publique.....	11
Avant-propos .....	13
Introduction générale .....	15

## Partie introductive

<b>Chapitre 1</b> L'évaluation dans un monde en constante évolution.....	21
<i>Jean Serge Quesnel, Vénétia Sauvain, Lynda Rey</i>	
<b>Chapitre 2</b> La terminologie de l'évaluation .....	37
<i>Richard Marceau et Francine Sylvain</i>	
<b>Chapitre 3</b> Les objets d'évaluation en contexte de développement international .....	57
<i>Jean Serge Quesnel et Marie-Hélène Adrien</i>	
<b>Chapitre 4</b> L'identification des parties prenantes en évaluation .....	73
<i>Kris Essosola Eale et Mary Pat Selvaggio</i>	
<b>Chapitre 5</b> La gestion d'une évaluation .....	85
<i>Jean Serge Quesnel et Ada Ocampo</i>	
<b>Chapitre 6</b> Principes, normes, règles et standards, éthique et déontologie de l'évaluation ....	101
<i>Jean Serge Quesnel et Ahmed Bencheikh</i>	
Le suivi et l'évaluation des objectifs de développement durable (ODD) : état des lieux et défis.....	119
<i>Mahmoud Ghoul</i>	
Enjeux et principes de l'évaluation des mouvements et politiques migratoires.....	127
<i>Diana Cartier et Christophe Franzetti</i>	
Évaluation des projets de coopération internationale sur l'adaptation aux changements climatiques : enjeux et défis.....	133
<i>Nicolas Montibert</i>	
Genèse et perspectives d'un mouvement africain de l'évaluation .....	141
<i>Oumoul Khayri Ba Tall, Carla-Anide Guillaume et Vénétia Sauvain</i>	
L'évaluation des politiques publiques en Afrique : pratiques, contraintes et perspectives.....	147
<i>Bobo B. Kabungu</i>	

## Partie I – Phase conceptuelle

<b>Chapitre 7</b>	<b>Le choix des approches évaluatives</b> .....	155
	<i>Thomas Delahais</i>	
	<b>Les approches économiques en évaluation : un aperçu</b> .....	181
	<i>Eva Anstett et Miché Ouédraogo</i>	
	<b>Value for money : Apprécier la valeur d’une intervention basée sur la performance de ses investissements</b> .....	185
	<i>Lynda Rey et Alexandre Fortin</i>	
<b>Chapitre 8</b>	<b>La modélisation des interventions en évaluation</b> .....	195
	<i>Juan Manuel Lopez Mejia et Lynda Rey</i>	
<b>Chapitre 9</b>	<b>Le processus d’évaluation</b> .....	215
	<i>Richard Marceau</i>	
<b>Chapitre 10</b>	<b>Les questions d’évaluation</b> .....	231
	<i>Richard Marceau</i>	

## Partie II – Phase méthodologique

<b>Chapitre 11</b>	<b>La matrice d’évaluation</b> .....	247
	<i>Richard Marceau</i>	
<b>Chapitre 12</b>	<b>Choix des outils pour les stratégies de collecte et d’analyse des données en évaluation</b> .....	261
	<i>Vénétia Sauvain, Ousséni Kinda et Maxime Desfossés</i>	
<b>Chapitre 13</b>	<b>Les méthodes qualitatives et quantitatives pour évaluer les effets</b> .....	287
	<i>Eva Anstett et Aref Ben Abdallah</i>	
	<b>Les devis expérimentaux pour évaluer les mécanismes de lutte contre la pauvreté : défis et enjeux</b> .....	311
	<i>Deo-Gracias Houndolo</i>	
	<b>L’engagement des enfants et des jeunes dans les évaluations</b> .....	319
	<i>Michele Tarsilla</i>	
	<b>De l’usage des technologies de l’information et de la communication en évaluation</b> .....	329
	<i>Tigawendé Prosper Kaboré et Faya François Bourouno</i>	

## Partie III – Phase empirique et analytique

<b>Chapitre 14</b>	<b>L'implication des parties prenantes en évaluation</b> .....	<b>339</b>
	<i>Kris Essosola Eale et Mary Pat Selvaggio</i>	
	<b>Participation des communautés locales, facteur de succès dans les évaluations adaptées à l'Afrique</b> .....	<b>359</b>
	<i>Anna Gueye et N'Diamé Gueye</i>	
<b>Chapitre 15</b>	<b>L'entrée sur le terrain en évaluation : enjeux et pratiques des collectes de données</b> .....	<b>365</b>
	<i>Vénétia Sauvain, Papa Souleymane Koulibaly, Nebnoma Kader Arnaud Sawadogo et Maxime Desfossés</i>	
<b>Chapitre 16</b>	<b>L'analyse des données pour répondre aux questions d'évaluation</b> .....	<b>383</b>
	<i>Miché Ouédraogo, Isabelle Brunet, Wend-Yam Saturnin Zoetyande, Maxime Desfossés et Vénétia Sauvain</i>	

## Partie IV – Phase de diffusion et de transfert

<b>Chapitre 17</b>	<b>La présentation des résultats d'évaluation</b> .....	<b>403</b>
	<i>Abdoul Aziz Sall et Élodie C. Roy</i>	
<b>Chapitre 18</b>	<b>L'élaboration des recommandations pour favoriser l'action</b> .....	<b>429</b>
	<i>Sylvie Lefrançois, Serge Eric Yakeu Djiam</i>	
<b>Chapitre 19</b>	<b>Un processus de transfert de connaissances pour favoriser l'utilisation des résultats d'évaluation</b> .....	<b>441</b>
	<i>Esther Mc Sween-Cadieux, Julie Lane, Kadidiatou Kadio, Valéry Ridde et Christian Dagenais</i>	

## Conclusion

<b>Futur de l'évaluation : les points de vue des premiers coordonnateur et présidents du Réseau Francophone de l'Évaluation</b> .....	<b>465</b>
<i>Guy Cauquil, Thierry Tsou Fematouo et Jean Serge Quesnel</i>	
<b>Conclusion</b> .....	<b>473</b>
<b>Notices biographiques des auteurs et autrices</b> .....	<b>477</b>





# Remerciement

Nous remercions le Gouvernement du Canada qui a financé cette publication, par l'entremise d'Affaires mondiales Canada, dans le cadre du projet intitulé Programme international de formation en évaluation du développement (PIFED).

Nous souhaitons particulièrement remercier nos pairs qui ont accepté de réviser et de commenter les chapitres pédagogiques : Jacques Bérard, David Buetti, Ludovic Décoret, Benoît Gauthier, Oumar Ndao Gning, Serguei Kimvi Mubian, Jean-François Lévesque, Cédric Ménard, Florence Morestin, Astou Niang, Joanes Hippolyte Sylli, Emmanuel Trépanier, Sara Vaca, Jackie Yiptong; ainsi que les quatre comités de *pifédiens* et *pifédiennes* qui ont jugé de la pertinence des chapitres réflexifs : Moustapha Abdoulaye G. Boulamine, Mulineca Calixte, Neslin Darbouze, Aubergine Flenon, Aimji Gofran, Abdoul Haidara, Meless Lathro, Joseph Lovely, Khaled Ahmed Naguib, Issoufou Nana, Soualiho Ouattara, Caroline Ouedraogo, Guéladio Aly Thiam, Sekou Traore, Elise Wang et Séraphin Yao.

Enfin, nos remerciements vont aux personnes de la Direction des affaires internationales qui ont contribué de près ou de loin à la publication de ce manuel.

Lynda, Jean Serge et Vénétia



# Mot du directeur général de l'École nationale d'administration publique

L'École nationale d'administration publique intervient depuis de nombreuses années dans le domaine de l'évaluation de programme et des politiques publiques, que ce soit sur le plan de la recherche, de l'offre de formations créditées et continues ou encore de l'accompagnement des administrations publiques au Québec, au Canada et à l'étranger, principalement en Afrique francophone et en Haïti. Ainsi, l'ENAP est fière d'avoir su mobiliser, parmi son réseau académique et de perfectionnement, 46 auteurs et autrices, de nombreux pairs réviseurs et des « pifédiens » et « pifédiennes » provenant du Québec, du Canada, de l'Europe, d'Afrique et d'Haïti pour vous offrir ce manuel pédagogique en langue française et en libre accès grâce à leur généreuse participation. Chacun et chacune représente une manière de réfléchir et d'agir avec et pour les actrices et acteurs présents dans les politiques publiques et la coopération internationale. L'ENAP leur adresse ses remerciements, ainsi qu'à toutes les personnes qui ont rendu possible la réalisation de ce manuel. La publication de cet ouvrage a été financée par le gouvernement du Canada, par l'entremise d'Affaires mondiales Canada, dans le cadre du Programme international de formation en évaluation du développement (PIFED). Ce projet, qui inclut différentes activités de renforcement des capacités en évaluation, est porté par l'ENAP depuis plus de 10 ans, et nous sommes fiers du chemin parcouru. Nous remercions Affaires mondiales Canada pour la confiance qu'elle accorde à l'ENAP dans le domaine de la coopération internationale depuis de nombreuses années.

Hugo Cyr, LL. D.





# Mot du directeur des affaires internationales de l'École nationale d'administration publique

Je suis fier que l'équipe de la direction des affaires internationales ait contribué à la publication de ce manuel tant attendu par la communauté des « pifédiens » et « pifédiennes » ! Je remercie tous les professeurs et professeures ainsi que les formateurs et formatrices du Programme international de formation en évaluation du développement (PIFED), les experts et expertes, les évaluateurs et évaluateuses et les nombreuses personnes qui ont accepté de participer à cette aventure sous divers chapeaux. Nous sommes particulièrement heureux de réunir dans cet ouvrage une diversité de perspectives qui sont au croisement de la théorie et de la pratique, de la coopération internationale et des politiques publiques, des modèles classiques et des modèles plus inclusifs, et qui représentent autant de pierres à l'édifice du renforcement des capacités en évaluation sur le plan international et en particulier dans la Francophonie.

Simon Chabot



# Avant-propos

*L'évaluation en contexte de développement : enjeux, approches et pratiques* est le fruit d'une collaboration sans précédent entre quarante-six personnes passionnées, praticiennes de l'évaluation de programmes. Corps professoral et étudiant, participants et participantes du Programme international de formation en évaluation du développement (PIFED), conseillers et conseillères ainsi que des consultantes et consultants œuvrant dans des pays d'Afrique, d'Haïti, d'Europe et d'Amérique du Nord, ont contribué de façon significative à la concrétisation de ce manuel collectif. Nous tenons ici à les remercier très chaleureusement pour leur générosité et le partage de leur expertise et de leur expérience. Enfin, nos remerciements vont à Affaires mondiales Canada, qui, par son appui financier, a rendu possible ce projet.

L'idée de cet ouvrage est née du constat que le champ de l'évaluation de programmes demeure largement dominé par des références anglophones, même si ces dernières années, on assiste au foisonnement d'ouvrages de référence francophones tels que *Approches et pratiques en évaluation de programmes* (Ridde et Dagenais, 2012); *L'évaluation : concepts et méthodes* (Brousselle et al., 2011); *Dictionnaire terminologique de l'évaluation : politiques, programmes, interventions. La dimension conceptuelle* (Marceau et Sylvain, 2014); *Évaluation des interventions de santé mondiale. Méthodes avancées* (Ridde et Dagenais, 2019); plus récemment, l'anthologie de textes fondateurs en évaluation, de source anglophone, *Évaluation. Fondements, controverses, perspectives* (Delahais et al., 2021). Ces derniers sont largement utilisés pour l'élaboration de cours de formation continue ou de cours universitaires, et ils sont inspirants pour la pratique évaluative. Par ailleurs, le manque de cas et d'expériences adaptés aux différents contextes des pays du Sud francophones (Afrique et Haïti principalement) face à la diversité de la production numérique anglophone demeure un défi auquel nous avons souhaité répondre, du moins partiellement.

Sans avoir la prétention d'être exhaustif, ce manuel contribue ainsi à combler un certain vide de littérature francophone sur l'évaluation en contexte de développement, c'est-à-dire le champ de l'évaluation investi par des bailleurs de fonds qui structurent la coopération internationale sous l'angle de l'aide au développement. *Il s'agit d'un contexte* particulier, à la fois politique et technique, qui implique des acteurs étatiques en partenariat avec des organisations internationales et des intermédiaires au profit de populations hétérogènes. *Il s'agit d'un contexte* où se juxtaposent à la fois les exigences de la gestion axée sur les résultats pour la conception et la mise en œuvre de grands projets de coopération internationale, et les contraintes propres au cycle de vie des politiques publiques et à la gouvernance des États. *Il s'agit d'un contexte* dans lequel cohabitent des pratiques et cultures organisationnelles multiples, parfois contradictoires. Ainsi, sans être exclusivement ancré dans le champ de la coopération internationale ni dans celui des politiques publiques, ce manuel tente d'apporter des éléments de contextualisation de la pratique évaluative dans un environnement complexe.

Répondant à des finalités essentiellement pédagogiques, il alterne entre des *chapitres généralistes*, présentant la vision de leurs auteurs et autrices sur la thématique couverte, indépendamment d'un contexte particulier, mais aussi des chapitres profondément ancrés dans leur environnement de pratique. Malgré la volonté d'une certaine cohérence pédagogique, la diversité des points de vue dans cet ouvrage est le reflet de la diversité des perspectives théoriques et pratiques de leurs auteurs et autrices.

Enfin, cet ouvrage s'inscrit également dans le cadre du PIFED offert par l'École nationale d'administration publique du Québec. Ce programme comprend plusieurs livrables, dont une formation, qui existe depuis 2011, et qui a évolué au gré des personnes qui en assurent la coordination pédagogique. Alors que la version anglophone du PIFED bénéficie de son manuel de référence, *The Road to Results : Designing and Conducting Effective Development Evaluations* (Morra Imas et Rist, 2009), il est temps de pouvoir offrir aux quelques 2000 « pifédiens » et « pifédiennes » quelque chose qui leur ressemble. Grâce à la contribution financière d'Affaires mondiales Canada, nous leur proposons, ainsi qu'à leurs pairs, principalement en Afrique et Haïti, cet ouvrage en libre accès. En plus d'offrir un socle conceptuel et pratique, ce manuel s'inscrit dans le mouvement international du renforcement des capacités en évaluation et espère ainsi contribuer à l'avancée de la culture évaluative.

## Références

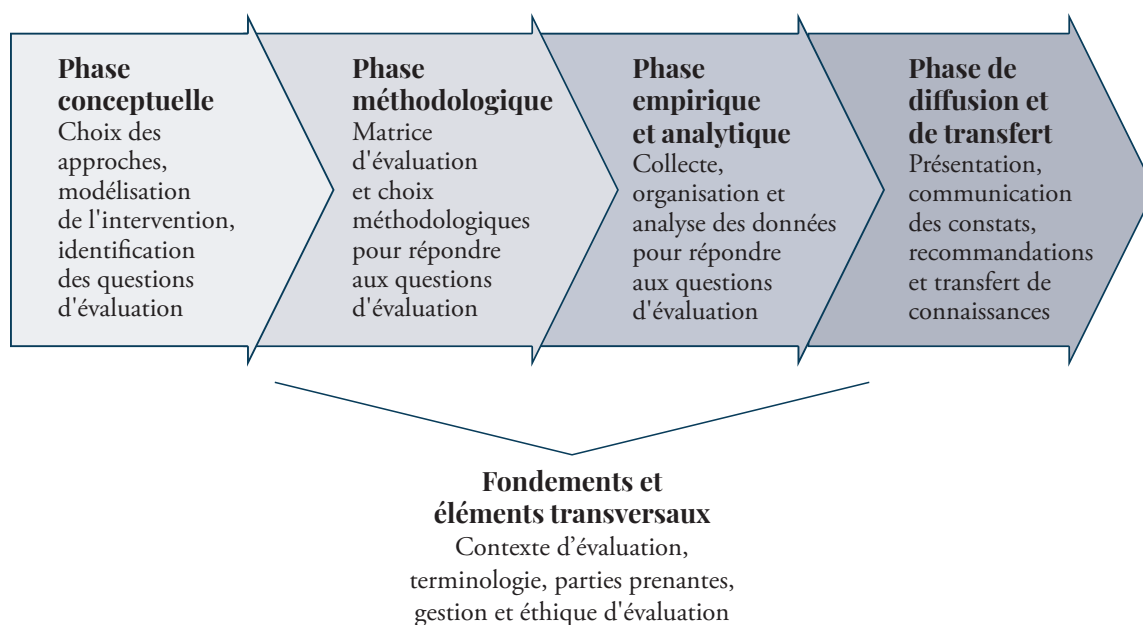
- Brousselle, A., Champagne, F., Contandriopoulos, A.-P. et Hartz, Z. (dir.) (2011). *L'évaluation : concepts et méthodes* (2<sup>e</sup> édition). Les Presses de l'Université de Montréal.
- Delahais, T., Devaux-Spatarakis, A., Revillard, A. et Ridde, V. (dir.) (2021). *Évaluation. Fondements, controverses, perspectives*. Éditions science et bien commun.
- Marceau, R. et Sylvain, F. (2014). *Dictionnaire terminologique de l'évaluation : politiques, programmes, interventions. La dimension conceptuelle*. Les Éditions GID.
- Morra Imas, L. G. et Rist, R. C. (2009). *The road to results: Designing and conducting effective development evaluations*. Banque mondiale.
- Ridde, V. et Dagenais, C. (dir.) (2012). *Approches et pratiques en évaluation de programmes*. Les Presses de l'Université de Montréal.
- Ridde, V. et Dagenais, C. (dir.) (2019). *Évaluation des interventions de santé mondiale. Méthodes avancées*. Éditions science et bien commun. IRD Éditions.



# Introduction générale

L'évaluation de programmes dans l'espace francophone, particulièrement en Afrique et en Haïti, connaît une progression sans précédent depuis une décennie. Bien que la demande demeure relativement faible, notamment au niveau des administrations publiques, à l'inverse des bailleurs de fonds internationaux qui en font une exigence, l'offre et l'intérêt pour ce champ sont en nette progression, comme l'attestent la création de sociétés nationales et l'émergence de professionnels en évaluation dans ces régions. En contribuant à la littérature francophone, cet ouvrage souhaite accompagner ce mouvement vertueux.

Ce manuel est structuré principalement autour de chapitres pédagogiques ayant pour objectif de présenter des concepts clés, des enjeux méthodologiques et des exemples issus de l'expérience des auteurs et autrices, mais également autour de chapitres réflexifs qui portent sur des tendances, des problématiques ou des approches particulières. Les différentes thématiques abordées sont réparties en suivant une démarche évaluative classique, fortement inspirée des processus de recherche (Fortin et Gagnon, 2016), c'est-à-dire : (1) une phase conceptuelle; (2) une phase méthodologique; (3) une phase empirique et analytique; (4) une phase de diffusion et de transfert, le tout étant introduit par une section sur les éléments communs à l'ensemble de la démarche évaluative.



**Figure 1**

Les phases de la démarche évaluative

Source : adapté de Fortin et Gagnon, 2016.

**Dans la partie introductive**, le premier chapitre présente des définitions clés de l'évaluation, son évolution dans un contexte d'incertitude et de bouleversements actuels profonds. Le chapitre sur la terminologie de l'évaluation est le produit de la réflexion des auteurs sur le sujet. Il structure de manière systémique l'ensemble du processus réflexif et conceptuel de l'évaluation. Il a le mérite de présenter un socle commun de concepts qui encadre ce processus. Dans le contexte du développement international, un chapitre est consacré spécifiquement aux objets d'évaluation du développement, qui prend en compte les évolutions de la coopération internationale en la matière. Un chapitre est consacré à la diversité des parties prenantes concernées et impliquées dans la conception d'une intervention et, a fortiori, de son évaluation : il insiste sur l'importance de les identifier dans la démarche évaluative. La démarche évaluative et les principes de gestion qui la structurent font également l'objet d'un chapitre à part entière, tout comme les principes, normes, règles et standards, qui façonnent l'éthique et la déontologie de la démarche. Plusieurs thèmes essentiels à l'évaluation en contexte de développement concluent cette première partie : un aperçu du contexte de suivi et d'évaluation des objectifs de développement durable (ODD), les enjeux et principes de l'évaluation des mouvements et politiques migratoires, les perspectives du mouvement africain de l'évaluation (*Made in Africa*), les caractéristiques de l'évaluation des politiques publiques en Afrique ou encore l'évaluation des projets de coopération internationale sur l'adaptation aux changements climatiques.

**La partie sur la phase conceptuelle** débute par un portrait dense des différentes familles d'approches évaluatives, en prenant le soin de définir la notion, et d'offrir ainsi un vaste choix de combinaisons possibles pour la pratique de l'évaluation. Une présentation succincte des approches économiques d'évaluation existantes et un regard particulier sur l'approche basée sur la valeur de l'investissement pour une intervention de King (*Value for money*) ont pour objectif de donner le goût au public de s'intéresser à l'évaluation du rendement. En effet, dans une perspective de développement durable et dans un contexte de ressources limitées, ces approches demeurent structurantes pour le rôle de l'évaluation dans la prise de décision publique. Quelle que soit la combinaison d'approches choisie, le lecteur ou la lectrice devra structurer, adapter et argumenter ses choix. Pour ce faire, la modélisation de l'intervention et la mobilisation d'un raisonnement structuré autour de grandes questions d'évaluation concernant l'intervention (appelé dans ce contexte « le processus d'évaluation ») demeurent des outils conceptuels puissants.

**La partie qui structure la phase méthodologique** vise à fournir un panorama non exhaustif des options méthodologiques pour mettre en œuvre l'évaluation, qui ont déjà fait l'objet de nombreuses publications francophones, d'autant plus que plusieurs d'entre elles sont également utilisées dans les processus traditionnels de recherche. Elle s'inscrit résolument dans des démarches rigoureuses de collecte de données issues des sciences économiques et sociales, tout en présentant quelques outils innovants, qui ont émergé dans le contexte du développement international, notamment pour résoudre des problématiques d'accès aux populations vulnérables (les enfants en particulier), ou encore pour mobiliser les technologies de l'information et de la communication pour la collecte et la visualisation des données.

**La phase empirique et la phase analytique**, habituellement séparées, ont été combinées pour des raisons pratiques et pour exprimer l'itération des stratégies de collecte et d'analyse de données pour répondre à des questions d'évaluation. Elle examine, d'une part, les différents enjeux de la pratique de la collecte de données à des fins d'évaluation, tels que l'implication concrète des parties prenantes aux différentes étapes de la démarche d'évaluation et les éléments de négociation de l'entrée sur le terrain. D'autre part, elle propose une synthèse des possibilités et des limites de l'analyse des données collectées pour formuler des constats et répondre aux questions d'évaluation.

Enfin, **la phase de diffusion et de transfert** clôt cet ouvrage, de telle sorte à accentuer la place des parties prenantes au cœur de la démarche évaluative, que ce soit au niveau de la présentation des livrables d'évaluation pour qu'ils soient à la fois compréhensibles et rigoureux pour leur public cible, de l'élaboration de recommandations solides, orientées vers l'action, mais réalistes, ou au niveau d'un processus de transfert de connaissances dynamique, imbriqué dans la démarche évaluative, et qui vise à favoriser l'utilisation des résultats d'évaluation pour une prise de décision éclairée.

## Références bibliographiques

Fortin, M.-F. et Gagnon, J. (2016). *Fondements et étapes du processus de recherche : Méthodes quantitatives et qualitatives* (3<sup>e</sup> édition). Chenelière Éducation.



# Partie introductive



# L'évaluation dans un monde en constante évolution



*Jean Serge Quesnel*

*Vénétia Sauvain*

*Lynda Rey*

## Introduction

Nous vivons dans un monde en constante évolution et sujet à des problématiques sociales complexes : pauvreté, changements climatiques, conflits, migrations, pandémies, etc. Aujourd'hui, plus que jamais, ces changements creusent les inégalités sociales intra et interpays, et rappellent la pertinence de « ne laisser personne de côté », en particulier les populations les plus vulnérables et marginalisées (Assemblée générale des Nations Unies, 2019). Par ailleurs, l'accès exponentiel aux nouvelles technologies de l'information et de la communication, en particulier aux réseaux sociaux, contribue à un foisonnement de données et de (dés)informations provenant de sources diversifiées aux dimensions géopolitiques. Pourtant, un fossé important demeure entre les connaissances scientifiques disponibles et leur utilisation dans la planification et la mise en œuvre d'une intervention. La contribution au développement, au changement social et au bien-être des individus, des organisations et de la société nécessite de reconnaître et d'utiliser des informations pertinentes, crédibles et robustes respectant des considérations éthiques, et d'y avoir accès. Une réflexion inédite a été menée lors de la publication de l'ouvrage de Goldman et Pabari sur la prise de décision basée sur des données probantes dans les politiques publiques dans l'espace africain, dont l'évaluation est l'une des possibilités qui s'offrent aux décideurs (Goldman et Pabari, 2020). Or, vient de paraître une anthologie de l'évaluation qui pose la question « à quoi sert l'évaluation ? » et dresse le bilan, plutôt faible, de l'influence de l'évaluation sur l'action publique (Delahais *et al.* 2021). Depuis ses débuts en 2014, l'un des axes du Forum International Francophone de l'Évaluation (FIFE) organisé par le Réseau Francophone de l'Évaluation (RFE) propose de débattre autour du dialogue entre la demande et l'offre d'évaluation. Dans le domaine de la coopération internationale, ce questionnement n'est pas nouveau : l'Agence française de développement (AFD) notait déjà avec justesse, en 2012, que si l'efficacité de l'aide au développement est régulièrement décriée, elle est sans doute la politique publique la plus évaluée (AFD, 2012).

Dans un monde en constante évolution, le rôle que peut jouer l'évaluation est-il toujours d'actualité ?

Dans ce chapitre, nous proposons quelques pistes de réflexion à cette vaste question, qui mériterait un ouvrage entier, en brossant un portrait synthétique, dans une perspective purement pédagogique, de ce qu'est l'évaluation, de son lien avec la notion de données probantes ou encore avec les autres fonctions issues de la démarche de performance et contribuant à la valeur publique; ainsi que de l'évolution des attentes et des finalités que l'on a pu avoir à son égard; et que l'on a désormais.

## Un monde en constante évolution : quelques éléments d'actualité

L'évaluation en tant que discipline et pratique s'inscrit dans un monde en constante évolution qui s'articule autour de plusieurs constats.

### Des enjeux sociaux complexes qui transcendent les frontières, les États et les sociétés

Le terme *problèmes pernicioeux* (*wicked problems*) fait son chemin dans la littérature<sup>1</sup> depuis quelques années pour caractériser des problèmes complexes, c'est-à-dire mal définis, dont les solutions, caractérisées par l'incertitude, sont ambivalentes et controversées, qui mettent en scène plusieurs niveaux d'interdépendance et des dynamiques sociales complexes (Alford et Head, 2017; Termeer *et al.*, 2015). Du niveau local à l'échelle planétaire, l'analyse de ces problèmes et leurs solutions potentielles ont des répercussions au-delà des contours géographiques, contextuels et sociaux qui les définissent. Par ailleurs, la réactivité et la mobilisation citoyennes dans un monde moderne, où la communication est extrêmement rapide, ont tendance à renforcer la politisation et l'interdépendance de ces problèmes (Head, 2018). C'est aussi le constat que font plusieurs auteurs du domaine de l'évaluation tel Patton, qui a proposé l'approche *Blue Marble Evaluation* à travers laquelle il énonce une série de principes à mettre en œuvre pour que l'évaluation, en tant que profession pluridisciplinaire, contribue à penser les problèmes dans leur globalité à l'échelle de la planète (Patton, 2020). Dans la même veine, un collectif d'auteurs a apporté des réflexions croisées sur le rôle de l'évaluation pour soutenir le changement et la transformation (Van den Berg *et al.*, 2021) à travers notamment : l'évaluation dynamique, l'évaluation basée sur les valeurs, ou encore l'importance de renforcer les systèmes nationaux pour soutenir les objectifs de l'Agenda 2030 des Nations Unies, ce dernier aspect ayant aussi fait l'objet d'un ouvrage récent de Persaud et Dagher (2021). Les auteurs analysent l'Agenda 2030 et son approche globale et multipartite pour parvenir à un processus de développement plus respectueux des droits de l'Homme et soucieux de l'environnement, et plaident en faveur de l'importance du suivi et de l'évaluation dans cet agenda. Ils soulignent les défis auxquels les évaluatrices et évaluateurs seront confrontés, autant en raison des caractéristiques inhérentes à l'Agenda (partenariats, différents niveaux de capacités des États, diversité d'acteurs, enjeux interconnectés, etc.), que de celles associées à la pratique et à la culture limitées de l'évaluation en général.

### Un foisonnement de données et d'informations à l'ère des « fausses nouvelles »

Depuis au moins une décennie, la quantité et la diversité des données générées par l'utilisation des solutions numériques et des services et applications associés, qui s'accumulent entre autres sous la forme de « mégadonnées », ainsi que la capacité de stockage et de traitement de celles-ci, révolutionnent le paysage de la connaissance scientifique, de la science du comportement et de la modélisation (Choumert-Nkolo et Phélinas, 2018). Malgré leur richesse et leur accumulation, ces données ne sont pas nécessairement utilisées (O'Hara, 2020). En parallèle, il devient facile pour différents acteurs et actrices malintentionnés de « manipuler » l'information et de répandre massivement et rapidement des mensonges, comme l'a tristement démontré le documentaire « *The Social Dilemma* » (Orlowski, 2020). Par ailleurs, la culture des *hashtags* et de l'information-réaction contribuent autant à lancer des mouvements sociaux de justice sociale de grande ampleur qu'à amplifier des confrontations

1. Que ce soit dans les domaines des politiques publiques, de l'évaluation de programme ou de la gestion.



inciviques et parfois contre-productives (Bouvier et Machin, 2021) ou encore à répandre de fausses informations à des fins politiques (Bovet et Makse, 2019). Dans ces nouveaux contextes, des défis de structuration, d'architecture et de réflexion sur l'utilisation de l'information et ses conséquences demeurent à l'échelle mondiale.

### **Un fossé important entre les connaissances scientifiques disponibles et leur utilisation**

Il existe un mouvement de plus en plus consensuel au sein de la communauté des évaluateurs et évaluatrices, selon lequel, pour favoriser son utilité, l'évaluation doit, plus que jamais, contribuer à la possibilité d'accéder à des informations pertinentes, crédibles et robustes, en somme des données probantes, et de les utiliser; et ceci, à l'échelle de la société et pas seulement au niveau des décideurs et décideuses.

« Une donnée probante est un fait ou une donnée utilisée, ou qui pourrait être utilisée, afin de prendre une décision, prononcer un jugement ou résoudre un problème » (Butcher, 1998, cité par McQueen et Anderson, 1999, p.80). Dans tous les cas, les données probantes indiquent si une proposition ou un constat sont vrais et valides. L'évaluation est une source de données probantes au même titre que les données issues de la recherche, des statistiques nationales, des données descriptives, des revues de littérature, de la mesure du rendement, du suivi d'une intervention, etc. À travers cette notion, « il s'agit de fonder la prise de décision politique et la pratique sur des données scientifiques les plus fiables » (Maunier, 2019, p.72). Or, l'enjeu de la prise de décision est essentiel dans ce concept, surtout utilisé dans le domaine biomédical, et les données probantes ne sont qu'un des éléments de prise de décision parmi d'autres tels que : les droits de la personne, le coût d'opportunité, l'adéquation culturelle, etc. (Archibald, 2015).

La bonne prise de décisions politiques dépend en grande partie de la qualité des informations fournies aux décideurs en temps opportun. Cette demande faite aux décideurs d'utiliser des données probantes pour la prise de décision repose sur le principe de responsabilisation démocratique (Head, 2016). Elle peut également venir de différents ministères, qui souhaitent réaliser au mieux un engagement, de comités parlementaires qui cherchent à examiner les progrès réalisés à la suite d'un plan ou à réviser leur budget, etc. (Goldman et Pabari, 2020). Dans ce contexte, l'utilisation de données probantes en politique permet une certaine transparence, réduit l'incertitude autour de la prise de décision, clarifie l'imputabilité et ainsi, peut servir d'outil contre le populisme. La littérature sur les données probantes et l'utilisation de celles-ci dans la conception de politiques publiques notamment, stipule que la collaboration transfrontalière, donc entre les évaluateurs et évaluatrices, la science et les décideurs politiques, est très importante (Danken *et al.*, 2016). Les approches basées sur la science fournissent une compréhension profonde de problèmes sociaux complexes, et la coconstruction des politiques publiques permettrait de mieux prendre en compte les multiples facettes de ceux-ci (Head, 2018). D'où l'importance de trouver un équilibre entre les données « scientifiquement valides » et « socialement légitimes » dans le processus de prise de décision ou de conception de politiques publiques (Brouselle *et al.*, 2011).

Néanmoins, la politique fait face à plusieurs enjeux, entre autres la gestion des risques politiques et des médias, leur volonté de conserver leur électorat, etc., ce qui limite l'utilisation de données probantes (Head, 2016). À l'inverse, l'apparition de problèmes sérieux (vécus comme des « crises politiques »), de polémiques ou de controverses publiques peut faire naître le besoin, pour les décideurs et les décideuses, d'obtenir des données probantes (Jacob et Rothmayr, 2012). Au-delà de la prise de décision politique, même s'il est attendu que les administrations publiques puissent proposer des

scénarios de prise de décision basés sur des données probantes, elles ne semblent pas être en mesure de jouer ce rôle. Les chiffres rapportés par Goldman et Pabari pour étayer ces propos sont édifiants : alors que 45 % des gestionnaires seniors de l'administration publique sud-africaine souhaiteraient utiliser des données probantes pour leur prise de décision, seulement 9 % indiquent être en mesure de le faire réellement (Paine Cronin et Sadan, 2015); ou encore, 85 % de la recherche scientifique dans le domaine de la santé n'est pas utilisée au niveau mondial (Stewart *et al.*, 2019), et seulement 1 % des investissements publics sous l'administration Obama étaient basés sur des données probantes (Bridgeland et Orszag, 2015). L'étude menée pour appuyer leurs propos indique que 40 % à 50 % des gestionnaires au Bénin, en Ouganda et en Afrique du Sud (368 gestionnaires pour cinq départements nationaux par pays) utilisent rarement ou jamais des données issues d'évaluation (Goldman *et al.*, 2020). Goldman et Pabari proposent un cadre conceptuel pour expliquer ce déficit à l'aide de la mécanique de l'offre et de la demande de données probantes. Afin d'être utilisées, les données devraient être accessibles, pertinentes et proposées en temps opportun (l'offre). En même temps, du point de vue de la demande, cette utilisation dépend de la volonté d'apprentissage, des capacités des acteurs à pouvoir interpréter et utiliser ces données afin de résoudre des problèmes, ainsi que de la structuration d'un environnement favorable. Un ensemble d'acteurs doivent pouvoir jouer le rôle de courtier de connaissances<sup>2</sup>, et l'ensemble de l'appareil est influencé par les acteurs externes, dont la société civile, mais également les lobbys (Goldman et Pabari, 2020).

Pour mieux comprendre la contribution potentielle de l'évaluation à l'édification des données probantes, il est utile de présenter la diversité de ses définitions.

## La diversité des définitions de l'évaluation

Une évaluation des interventions de développement, solide et indépendante, fournit des informations sur ce qui fonctionne et ce qui ne fonctionne pas, et en donne les raisons et les circonstances. Ces connaissances contribuent à améliorer l'efficacité et l'impact de l'aide au développement et permettent de rendre compte des résultats obtenus (OCDE, 2002). Il s'agit d'un élément de définition de l'évaluation parmi une diversité de propositions, issues de plusieurs domaines (académiques, coopération internationale, administration publique), qui ont évolué au fil du temps, en fonction de la pratique et des contextes. Le tableau ci-dessous en présente quelques-unes.

---

2. L'expression est amplement définie et précisée dans le contexte de l'évaluation dans le chapitre « Un processus de transfert de connaissances pour favoriser l'utilisation des résultats d'évaluation ».

**Tableau 1 — Une diversité des définitions de l'évaluation**

Auteurs	Définitions
Scriven, 1991.	« L'évaluation renvoie à un processus visant à déterminer le mérite et la valeur de quelque chose. »
Weiss, 1998.	« Une appréciation systématique de l'opérationnel et des impacts d'un programme ou d'une politique, à l'aune de standards explicites, de façon à améliorer le programme ou la politique. »
OCDE, 2002, p. 21.	« Appréciation systématique et objective d'un projet, d'un programme ou d'une politique, en cours ou terminée, de sa conception, de sa mise en œuvre et de ses résultats [...]. Une évaluation devrait fournir des informations crédibles et utiles permettant d'intégrer les leçons de l'expérience dans le processus de décision des bénéficiaires et des bailleurs de fonds. Le terme « évaluation » désigne également un processus aussi systématique et objectif que possible par lequel on détermine la valeur et la portée d'une action de développement projetée, en cours ou achevée. »
Patton, 2008.	« L'évaluation du programme est la collecte systématique d'informations sur les activités, les caractéristiques et les résultats des programmes pour juger un programme, pour améliorer l'efficacité du programme et/ou informer les décisions concernant la programmation future. »
Champagne <i>et al.</i> , 2011, p. 53.	« Évaluer consiste fondamentalement à porter un jugement de valeur sur une intervention en mettant en œuvre un dispositif capable de fournir des informations scientifiquement valides et socialement légitimes sur cette intervention ou sur n'importe laquelle de ses composantes, l'objectif étant de faire en sorte que les différents acteurs concernés, dont les champs de jugement sont parfois différents, soient en mesure de prendre position sur l'intervention pour qu'ils puissent construire, individuellement ou collectivement, un jugement susceptible de se traduire en actions. »
Secrétariat du Conseil du Trésor du Canada, 2016.	« Les évaluations visent à recueillir et à analyser de manière systématique et neutre des données probantes pour évaluer si un programme, une initiative ou une politique fonctionne, de même que pour évaluer la raison pour laquelle il fonctionne et la manière dont il fonctionne, dans le but d'éclairer la prise de décision, l'amélioration, l'innovation et la responsabilisation. »
Groupe des Nations Unies pour l'évaluation, 2018.	« Une évaluation a pour objet d'apprécier, de manière aussi systématique et impartiale que possible, une activité, un projet, un programme, une stratégie, une politique, un sujet, un thème, un secteur, un domaine opérationnel ou une performance institutionnelle. Elle permet d'analyser le niveau de réussite, à la fois en termes de résultats escomptés et inattendus, en examinant la chaîne des résultats, les processus, les facteurs contextuels et le lien de causalité. Elle s'appuie pour cela sur des critères tels que la pertinence, l'efficacité, l'efficience, l'impact et la durabilité. Une évaluation doit fournir, à partir d'éléments démontrables, des renseignements crédibles et utiles, qui permettent d'intégrer en temps voulu les conclusions, recommandations et enseignements dans le processus décisionnel des organisations et des parties prenantes. »

Quelle que soit la définition proposée, on peut constater des thèmes récurrents :

- L'évaluation correspond à une démarche systématique, qui repose sur la collecte et l'analyse de données, selon des normes, des standards et des principes éprouvés ainsi que sur des méthodologies documentées;
- L'évaluation permet de porter un jugement. Ce jugement peut être poussé très loin dans certaines définitions et vise « la valeur », « le mérite » ou « l'utilité » d'une intervention pour la société; et notamment en vue de sa transformation;
- L'évaluation s'inscrit dans un contexte de prise de décision et s'articule autour d'une position d'équilibre entre les différentes perspectives des acteurs et actrices, leur volonté ou non d'apprentissage et d'utilisation des évaluations.

La multiplicité des utilisations et des usages de l'évaluation a récemment fait l'objet d'études (Patton, 2020; Devaux-Spatarakis *et al.*, 2021), ainsi que l'absence d'utilisation, intentionnelle ou non, ou encore les « usages détournés » tels que la sélection des résultats de l'évaluation en fonction des intérêts politiques (Rajkotia, 2018). Dans le contexte de l'utilisation des données probantes issues des évaluations, la source de financement de l'évaluation peut avoir un impact sur la diffusion et l'utilisation de ses résultats (Gray, 2020). Les résultats d'évaluation qui concordent plus avec les points de vue politiques des dirigeants ont plus de chance d'être pris en considération (Fillol *et al.*, 2020). L'évaluation, qui est distincte de la recherche (celle-ci a globalement comme objectif la construction de la connaissance et non la prise de décision), bien qu'elle lui emprunte ses cadres conceptuels, ses méthodes et ses outils, ne dispose pas du même temps que celle-ci pour construire son argumentaire (Rey *et al.*, 2014). Par exemple, l'une des erreurs possibles d'utilisation de l'évaluation par les politiques est de penser que leurs résultats sont généralisables (Parkhurst, 2017), alors que ceux-ci sont généralement limités à leurs objets d'évaluation. Il apparaît alors important de délimiter la fonction d'évaluation par rapport à ce qu'elle n'est pas, mais également ce à quoi elle contribue, entre autres, à la performance (c'est-à-dire le « degré d'atteinte des résultats visés par une organisation » Secrétariat du Conseil du trésor du Québec, 2013) et à la valeur publique (c'est-à-dire les « attentes collectives des citoyens à l'égard du gouvernement et des services publics », Twizeyima et Andersson, 2019).

## L'évaluation : une fonction d'appui à la performance et de structuration de la valeur publique

Dans notre pratique en renforcement des capacités, nous avons constaté, particulièrement auprès de la clientèle internationale africaine francophone et haïtienne, que l'évaluation est encore souvent comparée, voir confondue avec d'autres fonctions, telles que le contrôle, l'audit/la vérification ou encore le suivi de la performance. Par exemple, l'évaluation est fréquemment perçue comme une forme additionnelle de vérification ou encore comme une étude de faisabilité (*ex ante*), un suivi de progrès (*in situ*) ou encore un synonyme de recherches académiques et théoriques sur les impacts (*ex post*). Le tableau ci-dessous présente des définitions de ces trois fonctions.

Lorsque la performance était définie sous l'angle « de l'assurance à la conformité aux lois et règlements en vigueur, au contrôle des moyens et à déontologie du service public » (Maltais, 2012), elle fut critiquée pour ne s'attarder qu'aux moyens et non aux résultats. L'avènement du nouveau management

public, en posant un regard critique sur les administrations publiques bureaucratiques orientées sur les moyens et sur la rentabilité des investissements publics pour l'atteinte de résultats, a contribué à structurer davantage la fonction de suivi de la performance.

Dans ce contexte, le suivi de la performance (*monitoring*) a joué un rôle clé pour tenter de pallier le simple examen de conformité que représentent les exercices d'audit-contrôle-vérification pour favoriser une meilleure compréhension du fonctionnement de la gestion publique à des fins d'apprentissages en continu. L'idée de base du suivi de la performance est, pour les organisations, de prévoir et de solutionner des problèmes de mise en œuvre des interventions qui peuvent avoir des conséquences fâcheuses sur l'atteinte des objectifs ou résultats attendus. Plus globalement, elle est associée à la planification stratégique, qui vise plutôt l'atteinte de résultats à l'échelle d'une organisation en fonction de la vision et des missions que celle-ci se donne ou qu'on lui attribue. Selon notre pratique, le suivi de la performance passe par une meilleure compréhension de la notion de résultats, par la clarification des objectifs à atteindre, par la définition d'indicateurs de performance, par une capacité institutionnelle à préparer des rapports de progrès et de suivi (accompagnée d'une volonté des rectifier les écarts), par des incitatifs pour encourager l'atteinte des résultats, y compris la reconnaissance des succès, la valorisation de l'apprentissage, l'application des leçons apprises de l'échec et par le contrôle de la fiabilité du système de suivi et d'évaluation.

**Tableau 2 — Contrôle – Audit – Vérification et suivi de la performance : quelques définitions**

Fonction	Qu'est-ce que c'est ?
Contrôle	« Une procédure de mise en œuvre en vue d'assurer de l'exactitude d'un enregistrement, du bon fonctionnement d'une structure, d'un service ou d'un système. » (Tremblay, 2012a).
Audit/ Vérification	« Activité de contrôle en matière de qualité, exercée de façon objective et indépendante, et destinées [sic] à améliorer les opérations d'une organisation et à en accroître la valeur » (OCDE, 2002, p. 17).  « La vérification ou l'audit est un examen critique d'une opération, d'une activité particulière ou de la situation générale d'une entité, prenant généralement appui sur des normes et des techniques établies et recommandées par un organisme professionnel, réalisé au moyen d'études, de contrôles ou de vérifications des décisions et des actes de gestion ou de leur conformité aux lois, aux normes ou aux règles établies, et menant généralement l'auditeur à produire au terme de ses travaux une communication écrite dans laquelle il exprime une opinion, un avis, une conclusion, ou encore des recommandations ou des mesures à prendre » (Tremblay, 2012b).
Suivi ( <i>monitoring</i> )	« Processus continu de collecte systématique d'informations, selon des indicateurs choisis, pour fournir [...] des éléments sur les progrès réalisés, les objectifs atteints et l'utilisation des fonds alloués » (OCDE, 2002, p. 27).

L'audit/vérification porte sur les processus de bonne administration, tandis que l'évaluation sonde l'ensemble des composantes des interventions : leur raison d'être, le rationnel qui justifie le choix de l'action pour la résolution d'un problème, ses retombées et les changements qu'elles produisent pour et dans la société. Par conséquent, elles se complètent. Le tableau qui suit compare les caractéristiques des deux fonctions.

**Tableau 3 — Caractéristiques comparées des fonctions audit/vérification et évaluation**

	Audit/Vérification	Évaluation
<b>Objet</b>	Cadre de gestion et processus	Politiques et interventions
<b>Focalisation</b>	Mandat et cadre du contrôle	Valeur et mérite des interventions
<b>Analyse</b>	Effizienz des systèmes et procédures	Multicritères
<b>Méthodologie</b>	Vérification financière/opérationnelle	Approches et méthodes mixtes qui évoluent avec sa théorisation, ses pratiques et les contextes de mise en œuvre
<b>Disciplines</b>	Audit, ingénierie, management	Sciences économiques et sociales
<b>Produit</b>	Vérification des contrôles de gestion	Jugement des effets des interventions
<b>Conséquences</b>	Rectification et ajustement	Changement transformationnel

L'évaluation contribue également à la structuration de la notion de « valeur publique ». Celle-ci renvoie à la poursuite, par les gouvernements, d'objectifs politiques et sociaux autres que les gains économiques, tels que l'efficacité du service public, l'égalité de traitement des usagers, l'inclusion sociale, l'ouverture, la régénération de la communauté, le bien-être de la communauté et la responsabilité (Twizeyimana et Andersson, 2019).

Dans les sociétés nord-américaines, les bureaux de vérificateurs/contrôleurs généraux dirigent la fonction étatique d'audit. La plupart de ces bureaux rendent compte auprès des instances législatives. La fonction de l'évaluation est surtout du ressort de l'exécutif étatique, particulièrement celui qui contrôle les dépenses (par exemple, l'Inspection Générale d'État au Sénégal, le Secrétariat du Conseil du trésor au Québec, ou encore un ministère des Finances et du Budget). Elle détermine l'efficacité et l'efficacité des interventions publiques à la lumière des résultats espérés, en conformité avec les politiques gouvernementales. Parfois, des commissions sont mises en place pour juger plus fondamentalement de la pertinence des interventions. Ces commissions peuvent être mandatées par les législateurs ou par l'exécutif, par exemple le Bureau d'Évaluation des Politiques Publiques au Bénin. Cela dit, on constate que la fonction d'évaluation des impacts ne semble pas toujours faire l'objet d'une délimitation claire dans les institutions publiques africaines francophones avec lesquelles nous avons travaillé jusqu'à maintenant. Dans le contexte de la coopération internationale, les évaluations sont exigées par les différents bailleurs de fonds.

## Évolution du rôle de l'évaluation au cours du temps

Il est désormais habituel de présenter les évolutions du rôle de l'évaluation au cours du temps sous l'expression « génération » empruntée à Guba et Lincoln (Dubois *et al.*, 2011; Guba et Lincoln, 1989; Hogan, 2007). Le tableau qui suit résume les générations et périodes distinctes dans l'histoire de l'évaluation.

**Tableau 4** — Les générations d'évaluations en fonction des périodes. Adapté de Dubois *et al.*, 2011, Guba et Lincoln, 1989, Guskey, 1999, Hogan, 2007, Madaus et Stufflebeam, 2020.

Génération d'évaluation	Périodes	Principales caractéristiques
<b>I</b>	Réformisme (1800-1900) Efficience et testage (1900-1930)	Mesure
<b>II</b>	Âge de l'innocence (1930-1960)	Description
<b>III</b>	Expansion (1960-1973) Professionnalisation et institutionnalisation (1973-1990) Intégration (1983-2000)	Jugement
<b>IV</b>	Utilitaire (2000 à nos jours)	Négociation

La première génération d'évaluation est teintée du rationalisme du *siècle des Lumières*. Durant la première période réformiste, l'évaluation vise à savoir dans quelles mesures les transformations qui ont été mises en œuvre améliorent les conditions de vie. À cette période, l'évaluation vise de façon particulière l'éducation. On promeut une collecte systématique de données pour appuyer le processus de prise de décision. On introduit des tests standardisés pour évaluer directement la performance des étudiants et indirectement celle des enseignants et des écoles. De plus, il y a de nombreuses commissions d'enquête qui donnent l'occasion d'institutionnaliser l'activité évaluative exploitant les possibilités offertes par la statistique.

La seconde période de la première génération d'évaluation continue à mettre l'accent sur l'efficience et le testage. L'influence de la *gestion scientifique* en quête d'une production optimale valorise la standardisation, la quantification, la précision (Taylor, 1911; Fayol, 1916). Cette influence de gestion incite tous les secteurs d'activités à rechercher des moyens pour systématiser leurs opérations et maximiser l'utilité de leurs interventions. Durant cette période, les termes évaluation et mesure sont interchangeables.

La deuxième génération d'évaluation émane d'une période historique marquée par des guerres mondiales, une récession économique sans précédent et une ouverture à la collaboration internationale par l'entremise de la création des Nations Unies et de ses agences. L'évaluation passe du stade d'examen du rendement individuel de l'écolier ou de l'ouvrier, à l'évaluation de la performance des programmes publics. En effet, les États investissent énormément dans des programmes sociaux. Selon George F. Madaus et Daniel L. Stufflebeam (1983), cette période fut appelée celle de *l'âge de l'innocence* parce qu'on vivait une grande croissance générant un grand optimisme, et on investissait beaucoup sans préoccupations d'imputabilité pour les fonds dépensés. À cette période, on croyait dans le bien-fondé des programmes sociaux, dont la finalité en soi ne pouvait être que positive, les programmes sociaux étant associés au progrès. Dorénavant, l'évaluateur ou l'évaluatrice n'est plus seulement un technicien de la métrique, c'est une personne qui décrit les interventions et examine leurs rendements. Bref,



évaluation n'est plus synonyme de mesure. Elle s'articule autour de la notion de contribution des interventions à des changements dans des contextes donnés, et veille à rendre explicites ces mécanismes de changement<sup>3</sup>.

La troisième génération d'évaluation comprend trois périodes que sont l'expansion, la professionnalisation/institutionnalisation et l'intégration de l'évaluation. Durant les années 1960 et 1970, les pays industrialisés jouissent d'une croissance économique. Il y eut l'émergence de nouvelles préoccupations tel l'activisme social en faveur des droits civiques, du féminisme, de la lutte contre la pauvreté. Les États investissent massivement dans des programmes sociaux, mais les premières évaluations ne font pas la preuve de leur valeur ajoutée. Cela a provoqué une suspicion chez les décideurs. Les médias rapportaient des cas de fraudes et des cas de mauvaise gestion. Par conséquent, on introduit de nouveaux outils de gestion : la gestion par objectifs et la planification, programmation et budgétisation. On considère que l'évaluation doit porter un jugement sur les interventions publiques. L'évaluation connaît une expansion sans précédent, puisque la législation de nombreux programmes oblige à une évaluation si le programme doit être reconduit ou augmenté. Ainsi, l'évaluation acquiert une légitimité accrue.

Au cours des années 1970, comme l'accent était mis sur le rôle de l'évaluateur ou de l'évaluatrice comme juge, les programmes de formation en évaluation furent investis par le milieu universitaire et axés sur la méthodologie. Un ensemble de facteurs ont contribué à faire de l'évaluation un champ professionnel distinct. Plusieurs publications spécialisées en évaluation apparaissent. Des associations professionnelles sont créées. Celles-ci promeuvent l'adoption de normes et de codes d'éthique. De plus, il se produit une institutionnalisation de la fonction d'évaluation. Les États ont adopté des politiques qui clarifiaient les rôles des acteurs institutionnels encourageant l'utilisation d'évaluation aux fins de la bonne gouvernance publique. La plupart des États ont mis en place une fonction d'évaluation, la logeant au sein d'une entité indépendante tels le *U.S. Government Accountability Office* (GAO), le Bureau du vérificateur général du Canada, et la Cour des comptes en France. De nombreux États ont créé des entités d'évaluation au sein de leurs ministères et agences gouvernementales.

Forte d'une professionnalisation accrue et d'une institutionnalisation plus systématique, l'évaluation est devenue plus intégrée à la gestion de la *chose publique*. Dans un contexte de contraintes, la fonction d'évaluation joue un rôle important en identifiant le rendement des programmes servant à guider l'allocation de ressources rares. Ainsi, parmi les programmes concurrents, on saura identifier les plus efficaces pour répondre aux besoins prioritaires.

La quatrième génération de l'évaluation peut être qualifiée d'*utilitaire*, faute d'une concordance entre les deux périodes historiques utilisées comme références. La rétrospective de l'évolution de l'évaluation au cours des années démontre clairement qu'une des contributions de l'évaluation est son apport pour mesurer, pour porter des jugements sur le rendement des interventions, pour déterminer des améliorations et pour éclairer les allocations de ressources. Jusqu'ici, l'approche analytique repose sur un modèle positiviste suivant lequel des experts appliquent des méthodes scientifiquement rigoureuses. En 1989, Egon G. Guba et Yvonna S. Lincoln proposent une nouvelle approche à l'évaluation et exposent les limites de l'approche positiviste de l'évaluation. Leur proposition oblige à un changement de paradigme (Guba et Lincoln, 1989).

---

3. Voir en particulier les chapitres : « Le choix des approches évaluatives », « La modélisation de l'intervention » et « Des méthodes d'évaluation des effets ».



L'évaluation n'est pas uniquement un jugement normatif, mais elle est surtout un jugement habilitant. L'évaluation n'est pas limitée par des normes fixes. Elle s'élabore dans un processus évolutif, itératif et interactif auquel participent les parties prenantes. Ainsi, il en découle des approches collaboratives, démocratiques et axées sur l'utilisation. L'évaluation est un instrument de négociation entre les parties prenantes de l'évaluation. Par conséquent, on voit l'essor exponentiel de l'évaluation participative. En effet, l'information générée par l'évaluation réfère aux besoins des parties prenantes. Elle émerge de la situation vécue et est basée sur une forme de théorie dont l'application varie d'une situation à une autre. En d'autres termes, la forme et le déroulement de l'évaluation découlent d'un processus d'ajustements négociés tenant compte des points de vue des parties prenantes. C'est une démarche évolutive qui permet une appropriation par les bénéficiaires des constats de l'intervention évaluée. Elle donne voix à la *demande*, alors que l'approche antérieure était dominée par l'*offre*. Il faut souligner qu'il ne s'agit pas de démarches opposées, mais qui se marient en portant une appréciation de la façon dont l'offre est reçue par les demandeurs. En effet, ces démarches ne sont pas mutuellement exclusives, d'où les approches participatives et systémiques.

À la lumière de la Déclaration de Paris et du Programme d'action d'Accra, il est apparu clair que les gouvernements nationaux (davantage que les bailleurs de fonds) sont ultimement responsables et imputables de leurs propres actions. Force est de constater que les méthodes d'évaluation doivent évoluer pour juger les impacts des interventions de plus en plus complexes qui sont mises en œuvre dans des contextes pluridimensionnels impliquant de multiples partenaires qui doivent contribuer et avoir voix au chapitre en matière d'évaluation. La prise en compte des problématiques de l'égalité de genre, de l'équité, notamment envers les besoins des groupes vulnérables tels les enfants, les personnes âgées et les personnes handicapées, demeure un élément essentiel de cette complexité.

## **L'évaluation au 21<sup>e</sup> siècle; une nouvelle génération ? Celle de la transformation ?**

Les réflexions actuelles sur l'évolution de l'évaluation ne sont pas si récentes. Des tendances qui avaient été identifiées à l'aube du 21<sup>e</sup> siècle demeurent pertinentes (Blaine *et al.*, 2004) :

1. Une importance accordée à la légitimité et à la validité des évaluations conduites à l'interne par les organisations, sans nécessairement minimiser le rôle des évaluations externes;
2. Une utilisation accrue des méthodes mixtes, qualitatives et quantitatives, en fonction des contextes;
3. Le développement d'approches d'évaluations axées sur les changements à l'échelle d'une société, voire de la planète, qui sont qualifiées de transformations. Ces approches tiennent compte des différentes dimensions de la complexité des interventions dans une perspective systémique;
4. Une augmentation de la professionnalisation des évaluateurs et évaluatrices ainsi que des préoccupations envers les questions d'éthique dans la pratique des évaluations;
5. Une hausse importante de l'utilisation des évaluations par le monde des affaires et l'industrie, par les fondations et les agences du secteur non gouvernemental;
6. L'utilisation de l'évaluation pour habilitier les parties prenantes, et pour les impliquer dans toutes les étapes d'évaluation pour que celle-ci soit utile;

7. L'appropriation de nouveaux outils informatiques pour augmenter la capacité des évaluations à disposer de constats généralisables (en raison de la taille de la population étudiée);
8. Au sein des appareils étatiques, une décentralisation et une délégation des responsabilités de la fonction évaluative à tous les paliers.

## Conclusion

L'évaluation est devenue plus complexe, comme en témoignent ses évolutions, les ramifications de ses nombreuses utilités potentielles ou avérées et le foisonnement des approches qu'elle a générées. À nos yeux, elle demeure essentielle pour une prise de décision éclairée. Elle évolue pour tenir compte des contextes nationaux différents tout en restant mondialement interconnectée, pour appuyer une collaboration centrée sur une réponse aux besoins définis par les ressortissants nationaux, tout en respectant leurs choix de développement. Elle répond aux besoins de transparence, d'efficacité et de redevabilité dans le domaine complexe de l'aide au développement. Puisque la fonction d'évaluation est compréhensive et réflexive, elle joue un rôle important dans cet environnement de changement et d'amélioration continue. Aux dires de la Banque africaine de développement :

Une culture d'évaluation est généralement reconnue comme une solution clé pour améliorer l'efficacité et l'efficience des gouvernements, organisations de développement et institutions financières internationales. La documentation existante suggère que pour que les fonctions de gestion des performances, d'évaluation et de recherche soient efficaces, les organisations et les institutions ont besoin d'une solide culture d'évaluation – une culture où les preuves sont délibérément recherchées afin de mieux concevoir, mettre en œuvre, exécuter les interventions. (BIE-BAD, 2020)

Dans le contexte du développement international, dans un monde en mutation, le défi pour la fonction de l'évaluation durant le XXI<sup>e</sup> siècle est celui de « ne laisser personne de côté » (Assemblée générale des Nations Unies, 2019). En effet, comment extraire au moins 730 millions de personnes de l'extrême pauvreté et jusqu'à 2,3 milliards de personnes de la fragilité (Banque mondiale, 2018) grâce à des stratégies de développement plus efficaces, plus endogènes, plus équitables et plus durables ? L'évaluation doit examiner comment les politiques, les investissements et les partenariats devraient être articulés autour de l'objectif fondamental d'assurer un développement inclusif, équitable et durable. L'évaluation doit soutenir les objectifs de développement durable à l'horizon 2030 et les années subséquentes. Pour cela, il faut adopter des approches innovatrices et participatives, et recueillir des données désagrégées, de manière à déterminer quelles personnes sont les plus pauvres, les plus marginalisées et les plus vulnérables, définir les facteurs d'inégalité et d'exclusion, et proposer des mesures de solutions pragmatiques à la problématique. L'évaluation doit adopter une vision systémique du développement et des approches « adaptées à l'Afrique » dans une perspective de décolonisation des savoirs. C'est une transformation qui nécessite non seulement des politiques et des investissements, mais aussi une modification profonde des habitudes, voire un changement de paradigme. C'est un appel à créer une nouvelle génération d'évaluation, celle en référence à la planète bleue selon l'approche « *Blue Marble* » (Patton, 2020), basée sur des principes qui prennent en compte le caractère systémique des problèmes, leur interconnexion à l'échelle mondiale, et leurs conséquences dramatiques à l'échelle locale, pour susciter un changement transformationnel. L'objectif n'est pas moins que de structurer une pratique évaluative transdisciplinaire, transculturelle, transfrontalière et transformatrice.

## Références bibliographiques

- Agence Française de Développement/European Development Research Network (AFD/EUDN) (dir.). (2012). Remerciements [Actes de conférence]. Malaise dans l'évaluation : quelles leçons tirer de l'expérience de développement ? 9<sup>e</sup> conférence AFD/EUDN. *Revue d'économie du développement*, 20, 5-6. <https://doi.org/10.3917/edd.264.0005>
- Alford, J. et Head, B. W. (2017). Wicked and less wicked problems : a typology and a contingency framework. *Policy and Society*, 36(3), 397-413. <https://doi.org/10.1080/14494035.2017.1361634>
- Archibald, T. (2015). “They Just Know” : The epistemological politics of “evidence-based” non-formal education. *Evaluation and Program Planning*, 48, 137-148. <http://dx.doi.org/10.1016/j.evalprogplan.2014.08.001>
- Assemblée générale des Nations Unies. (2019, 9 septembre). *Déclaration politique issue du Sommet sur les objectifs de développement durable*, A/HLPF/2019/L.1. Forum politique de haut niveau pour le développement durable. <https://undocs.org/fr/A/HLPF/2019/L.1>
- Banque mondiale. (2018). *Poverty and Shared Prosperity 2018 : Piecing Together the Poverty Puzzle*. <https://openknowledge.worldbank.org/bitstream/handle/10986/30418/9781464813306.pdf>
- BIE-BAD. Bureau indépendant d'évaluation, Banque africaine de développement. (2020, mars). *Promouvoir une culture de l'évaluation en 2020 et au-delà*. eVALUation Matters.
- Blaine, R., Fitzpatrick, J. L. et Sanders, J. R. (2004). *Program Evaluation : Alternative Approaches and Practical Guidelines* (3rd ed.). Allyn & Bacon.
- Bouvier, G. et Machin, D. (2021). What gets lost in Twitter ‘cancel culture’ hashtags? Calling out racists reveals some limitations of social justice campaigns. *Discourse & Society*, 32(3), 307–327. <https://doi.org/10.1177/0957926520977215>
- Bovet, A. et Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature communications*, 10(1), 7. <https://doi.org/10.1038/s41467-018-07761-2>
- Bridgeland, J. et Orszag, P. (2013). Can government play money ball? *The Atlantic*, juillet/août. <https://www.theatlantic.com/magazine/archive/2013/07/can-government-play-moneyball/309389/>
- Brousselle, A., Champagne, F., Contandriopoulos, A.-P. et Hartz, Z. (dir.). (2011). *L'évaluation : concepts et méthodes* (2<sup>e</sup> éd.). Les Presses de l'Université de Montréal.
- Butcher, R.B. (1998). *Foundations for evidence-based decision making. Canada health action : building on the legacy, vol. 5 : Evidence and Information*. Éditions Multimondes.
- Champagne, F., Brousselle, A., Contandriopoulos, A.-P., Denis, J.-L. et Hartz, Z. (2011). L'évaluation dans le domaine de la santé : concepts et méthodes. Dans A. Brousselle, F. Champagne, A.-P. Contandriopoulos et Z. Hartz (dir.), *L'évaluation : concepts et méthodes* (2<sup>e</sup> éd.). Les Presses de l'Université de Montréal.
- Choumert-Nkolo, J. et Phélinas, P. (2018). Nouveaux principes d'élaboration des enquêtes ménage dans les pays du Sud. *Revue d'économie du développement*, 3(26), 121-151. <https://doi-org.acces.bibl.ulaval.ca/10.3917/edd.323.0121>

- Danken, T., Dribbisch, K. et Lange, A. (2016). Studying Wicked Problems Forty Years On : Towards a Synthesis of a Fragmented Debate. *Zeitschrift für Public Policy, Recht und Management*, 9(1), 15-33. <http://dx.doi.org/10.3224/dms.v9i1.23638>
- Delahais, T., Devaux-Spatarakis, A., Revillard, A. et Ridde, V. (dir.). (2021). *Évaluation – Fondements, controverses, Perspective*. Éditions science et bien commun.
- Dubois, C-A., Champagne, F. et Bilodeau, H. (2011). Historique de l'évaluation. Dans A. Brousselle, F. Champagne, A.-P. Contandriopoulos et Z. Hartz (dir.) *L'évaluation : concepts et méthodes* (2<sup>e</sup> éd.). Presses de l'Université de Montréal.
- Fayol, H. (1916). *Administration industrielle et générale*. Dunod.
- Fillol, A., Kadio, K. et Gautier L. (2020). L'utilisation des connaissances pour informer des politiques publiques : d'une prescription technocratique internationale à la réalité politique des terrains. *Revue française des affaires sociales*, 2020(4), 103-127. <https://doi-org.acces.bibl.ulaval.ca/10.3917/rfas.204.0103>
- Golman, I. et Pabari, M. (2020). *Using Evidence in Policy and Practice : Lessons from Africa*. Routledge.
- Gray, L. (2020). Evidence-based policy-making and exam board insider researchers : creating communicative spaces. *Assessment in Education : Principles, Policy & Practice*, 27(2), 142-159. <https://doi.org/10.1080/0969594X.2020.1749557>
- Groupe des Nations Unies pour l'évaluation (UNEG). (2018). *Normes et règles d'évaluation*. <http://www.unevaluation.org/2016-Norms-and-Standards>
- Guba, E. G. et Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. SAGE Publications.
- Guskey, T. R. (1999). *Evaluating Professional Development*. Corwin Press.
- Head, B. W. (2016). Toward More "Evidence-Informed" Policy Making? *Public Administration Review*, 76(3), 472-484. <https://doi.org/10.1111/puar.12475>
- Head, B.W. (2018). Forty years of wicked problems literature : forging closer links to policy studies. *Policy and Society*, 38(2), 180-197. <https://doi.org/10.1080/14494035.2018.1488797>
- Hogan, R.L. (2007). The Historical Development of Program Evaluation : Exploring the Past and Present. *Online Journal of Workforce Education and Development*, 2(4).
- Jacob, S. (2009). Opération chloroforme ou la réinvention de l'État rationnel : l'évaluation et les données probantes. *Criminologie*, 42(1), 201-223. <https://doi.org/10.7202/029813ar>
- Jacob, S. et Rothmayr, C. (2012). L'analyse des politiques publiques. Dans V. Ridde et C. Dagenais (dir.), *Approches et pratiques en évaluation de programmes* (2<sup>e</sup> éd.). Presses de l'Université de Montréal.
- Madaus, G. et Stufflebeam D. (2000). Program Evaluation : A Historical Overview. Dans D. L. Stufflebeam, G. F. Madaus et T. Kellaghan (dir.), *Evaluation Models. Evaluation in Education and Human Services*, vol 49. Springer, Dordrecht. [https://doi.org/10.1007/0-306-47559-6\\_1](https://doi.org/10.1007/0-306-47559-6_1)
- Maltais, D. (2012). Performance et gestion de la performance. Dans L. Côté et J.-F. Savard (dir.), *Le Dictionnaire encyclopédique de l'administration publique*. [www.dictionnaire.enap.ca](http://www.dictionnaire.enap.ca)
- Maunier, S. (2019). Données probantes : quel rôle pour la recherche qualitative? *Recherches qualitatives*, 38(1), 71-87. <https://doi.org/10.7202/1059648ar>

- Mc Donnell, I. et Malhotra, R. (2018). Synthèse : La coopération pour le développement à l'horizon 2030 – se renouveler et se réformer pour concrétiser l'engagement de ne laisser personne de côté. Dans OCDE. (dir.), *Coopération pour le développement 2018. Agir ensemble pour n'oublier personne*. <https://doi.org/10.1787/dcr-2018-fr>
- McQueen, D. et Anderson, L. M. (1999). Données probantes et évaluation des programmes en promotion de la santé. *Ruptures, revue transdisciplinaire en santé*, 7(1), 79-98.
- Ménard, L., Arsenault, M. et J.-F. Joly (2004). *Dictionnaire de la comptabilité et de la gestion financière*. Institut canadien des comptables agréés.
- OCDE. (2002). *Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats*. <https://www.oecd.org/dac/evaluation/2754804.pdf>
- O'Hara, K. (2020). The contradictions of digital modernity. *AI & Society*, 35, 197-208. <https://doi.org/10.1007/s00146-018-0843-7>
- Orlowski, J. (réalisateur) et Rhodes, L. (producteur). (2020, 9 septembre). *The Social Dilemma*.
- Paine Cronin, G. et Sadan, M. (2015). Use of evidence in policy making in South Africa : An exploratory study of attitudes of senior government officials. *African Evaluation Journal*, 3. <https://doi.org/10.4102/aej.v3i1.145>
- Parkhurst, J. (2017). *The politics of evidence : from evidence-based policy to the good governance of evidence*. Routledge.
- Patton, M. Q. (2020). *Blue Marble Evaluation : Premises and Principles*. The Guilford Press.
- Patton, M. Q. (2021). *Utilization-focused evaluation* (4<sup>e</sup> éd., 2008). SAGE Publications Inc.
- Persaud, N. et Dagher, R. (2021). *The Role of Monitoring and Evaluation in the UN 2030 SDGs Agenda*. Springer Link.
- Rajkotia, Y. (2018). Beware of the Success Cartel : A Plea for Rational Progress in Global Health. *BMJ Global Health*, 3(6). <https://doi.org/10.1136/bmjgh-2018-001197>
- Rey, L., Tremblay, M. C. et Brousselle, A. (2014). Managing tensions between evaluation and research : illustrative cases of developmental evaluation in the context of research. *American Journal of Evaluation*, 35(1), 45-60. <https://doi.org/10.1177/1098214013503698>
- Schalock, R. L., Gomez, L. E., Verdugo, A. M. et Claes, C. (2017). Evidence and Evidence-Based Practices : Are We There Yet? *Intellectual and Development Disabilities*, 55(2), 112-119. <https://doi.org/10.1352/1934-9556-55.2.112>
- Scriven, M. (1991). *Evaluation thesaurus*. SAGE Publications.
- Secrétariat du Conseil du Trésor du Canada. (2016). *Politique sur les résultats. Annexe A : Définitions*. <https://www.tbs-sct.gc.ca/pol/doc-fra.aspx?id=31300>
- Secrétariat du Conseil du trésor du Québec. (2013). *Glossaire des termes usuels en mesure de performance et en évaluation. Pour une gestion saine et performante*. [https://www.tresor.gouv.qc.ca/fileadmin/PDF/publications/glossaire\\_termes\\_usuels.pdf](https://www.tresor.gouv.qc.ca/fileadmin/PDF/publications/glossaire_termes_usuels.pdf)
- Stewart, R., Dayal, H., Langer, L. et van Rooyen, C. (2019). The evidence ecosystem in South Africa : Growing resilience and institutionalisation of evidence use. *Palgrave Communications*, 5, 1–12. <https://doi.org/10.1057/s41599-019-0303-0>

- Taylor, F. W. (1911). *The Principles of Scientific Management*. Harper and Bros.
- Termeer, J. A. M. C., Dewulf, A., Breeman, G. et Stiller, S. J. (2015). Governance Capabilities for Dealing Wisely With Wicked Problems. *Administration & Society*, 47(6), 680-710. <https://doi.org/10.1177%2F0095399712469195>
- Tremblay, M.-S. (2012a). Contrôle. Dans L. Côté et J.-F. Savard (dir.), *Le Dictionnaire encyclopédique de l'administration publique*. [www.dictionnaire.enap.ca](http://www.dictionnaire.enap.ca)
- Tremblay, M.-S. (2012b). Vérification/audit. Dans L. Côté et J.-F. Savard (dir.). *Le Dictionnaire encyclopédique de l'administration publique*. [www.dictionnaire.enap.ca](http://www.dictionnaire.enap.ca)
- Twizeyimana, J. D. et Andersson, A. (2019). The Politics of Open Government Data : Understanding Organizational Responses to Pressure for More Transparency. *Government Information Quarterly*, 36(2), 167-178.
- Van den Berg, R. D., Magro, C. et Adrien, M-H. (2021). *Transformational Evaluation for the Global Crises of our times*. IDEAS.
- Weiss, C. (1998). Theory-based evaluation : past, present and future. *New Directions for Evaluation*, 76, 41-55.



# La terminologie de l'évaluation

*Richard Marceau  
Francine Sylvain*



## Introduction

Les mots ne coûtent pas cher. Ils voyagent à travers le temps et l'espace sans guère de restrictions. Ils parviennent de civilisations anciennes ou instantanément de l'autre bout de la planète. Chez soi, au bureau ou à la bibliothèque, on dispose de dictionnaires de langue générale qui contiennent des centaines de milliers de mots. Des lexiques et des glossaires, plus spécialisés, sont offerts gratuitement et en grand nombre sur le Web.

Les évaluateurs et les évaluatrices d'aujourd'hui peuvent puiser des mots dans de nombreuses sources pour communiquer avec les parties prenantes d'une intervention, ou avec leurs collègues. La gratuité et l'abondance de ces sources pourraient cependant nous amener à sous-estimer le rôle et la valeur des mots en évaluation.

Ce serait une erreur ! Car les mots appartenant à une science ou à une pratique professionnelle, appelés termes dans la discipline de la terminologie, portent en eux des concepts. Et ces concepts constituent la base à partir de laquelle les idées prennent vie, se développent et forment un raisonnement. Dans la discipline de l'évaluation, ces concepts donnent du sens à la pratique évaluative. Sans eux, l'évaluation serait un chaos intellectuel, et ne pourrait prétendre être un champ de connaissances et de pratique. La nécessité de cohérence, de pertinence et de communication efficace milite pour la rigueur dans le choix des mots et dans leur usage.

Le présent chapitre puise dans les recherches et dans les travaux que nous avons menés sur la terminologie de l'évaluation dans le but de formaliser davantage les fondements de l'évaluation. Le *Dictionnaire terminologique de l'évaluation*, paru en 2014, et le *Dictionnaire multilingue de l'évaluation* (DME), à paraître, en sont issus. Nous faisons ici un résumé de notre approche à l'égard de la terminologie en évaluation et présentons à la fin du chapitre un glossaire des concepts de base en évaluation contenant l'ensemble des termes du DME.

Plus précisément, 1) nous jetons un bref regard sur la formalisation dans les sciences en général; 2) nous nous attardons sur la relation qui existe entre la formalisation de l'évaluation et l'affirmation de son identité face aux disciplines connexes, et nous voyons comment la terminologie a un rôle à y jouer; 3) nous appliquons une démarche découlant de la théorie générale de la terminologie associée à Eugen Wüster (L'Homme, 2004) pour proposer une terminologie propre à l'évaluation; 4) nous voyons comment la terminologie de l'évaluation encadre et guide la pratique évaluative; 5) nous amorçons une discussion sur l'apport des différentes langues et cultures à la terminologie de l'évaluation, et en particulier de la francophonie. Enfin, nous proposons un glossaire de 103 termes propres au domaine de l'évaluation.

## La formalisation dans les sciences

Les disciplines scientifiques et les pratiques professionnelles formalisent leurs concepts de base, leurs idées maîtresses. La psychologie, la physique, la biologie, l'économique, par exemple, apportent de la rigueur à leur pensée et à leur pratique, améliorent leurs communications et assurent leur crédibilité par la formalisation de leurs concepts.

Ainsi, les concepts de la physique sont la plupart du temps représentés sous forme mathématique (ex. :  $E = mc^2$ ). La chimie formalise par une symbolique les transformations des composés de la matière (ex. :  $\text{NH}_3 + \text{O}_2 \rightarrow \text{NO} + \text{H}_2\text{O}$ ). La biologie dispose d'une nomenclature ordonnée (nom, famille, type) qui lui permet de classer la flore, tout comme la faune d'ailleurs, en exprimant rigoureusement les liens de parenté entre les espèces (ex. : Nom : *Juniperus*; Famille : Cupressacées; Type : Conifère).

Ces quelques exemples ne sont que des illustrations de la manière dont les disciplines scientifiques formalisent leurs idées maîtresses, mais suffisent à montrer que la formalisation des concepts contribue à clarifier et à préciser la pensée et la pratique tout en facilitant la communication. Cette formalisation passe inévitablement par la définition des concepts et par l'établissement de relations précises entre eux, que ce soit par des moyens mathématiques, logiques ou langagiers. En conséquence, il apparaît que la terminologie a un rôle naturel à jouer dans la formalisation d'une discipline.

## La formalisation de l'évaluation et l'affirmation de son identité

L'évaluation est une discipline encore jeune. Apparue au milieu du 20<sup>e</sup> siècle, elle se situe à un carrefour de plusieurs disciplines scientifiques et de pratiques professionnelles déjà bien affirmées. L'évaluation emprunte de nombreux pans méthodologiques et techniques à plusieurs disciplines des sciences sociales. Les statistiques, les méthodes d'enquête, les analyses avantages-coûts, par exemple, ont été développées par d'autres champs disciplinaires, et elles rendent ainsi un grand service à l'évaluation. Incidemment, ces méthodes et techniques portent en elles beaucoup de rigueur et de formalisme, ce qui contribue à donner de la crédibilité à l'évaluation en tant que discipline et pratique.

Mais, en contrepartie, ces emprunts soulèvent des questions sur l'identité propre de l'évaluation. Par exemple, quels sont les concepts « identitaires » de l'évaluation, ceux qui particularisent l'évaluation, ceux qui la distinguent des disciplines connexes? Ces concepts de base, ces idées maîtresses, sont-ils nombreux? Sont-ils rares? Sont-ils vagues ou précis, disparates ou reliés les uns aux autres? Les relations entre les concepts sont-elles floues ou formalisées? A-t-on tenté d'uniformiser les concepts et les termes utilisés dans les différentes communautés d'évaluation?

Il est malheureusement difficile de répondre à ces questions, car, à notre connaissance, peu d'efforts ont été consacrés jusqu'à présent à la formalisation des concepts de l'évaluation à l'échelle de la discipline. Pourtant, si l'évaluation n'offre pas un corpus substantiel, pertinent et cohérent de concepts de base qui lui sont propres, c'est l'identité même de la discipline qui est en jeu. Sans une pensée évaluative cohérente et une terminologie organisée, l'évaluation peine à être reconnue, à se distinguer des champs connexes comme l'audit, la science économique ou les sciences politiques. Nul doute selon nous, la discipline de l'évaluation a grand avantage à disposer d'une structure formelle de sa terminologie pour qu'elle puisse, comme elle le souhaite, contribuer pertinemment et significativement à l'amélioration du sort de nos sociétés.



## Application d'une démarche terminologique au domaine de l'évaluation

Dans la présente section, nous proposons une application de la théorie générale de la terminologie pour les termes de base du domaine de l'évaluation. Pour ce faire, nous devons :

- a. délimiter un ensemble de concepts de base;
- b. identifier les termes correspondants;
- c. créer des néologismes;
- d. classer les concepts dans un système conceptuel;
- e. rédiger des définitions.

### a. Délimitation des concepts de base

Nous avons élaboré une démarche de conceptualisation en cinq étapes qui nous conduit à construire ce que nous considérons être les grandes catégories de concepts de base de l'évaluation. Le tableau 1 présente notre démarche de conceptualisation de la terminologie de base de l'évaluation.

**Tableau 1 — Démarche de conceptualisation de la terminologie de base de l'évaluation**

Étape	Contenu des étapes
Étape 1	Définir ce qu'est l'évaluation.
Étape 2	Identifier les objets de l'évaluation.
Étape 3	Préciser les caractéristiques et les propriétés générales des objets de l'évaluation.
Étape 4	Préciser les éléments du processus d'évaluation.
Étape 5	Expliciter les critères d'évaluation.

Chacune des cinq étapes de cette démarche de conceptualisation produit un sous-ensemble de concepts de base partageant des caractéristiques communes. En découlent des concepts de base qui sont naturellement regroupés en cinq grandes catégories.

Le tableau 2 présente les cinq grandes catégories de concepts de base, dotées chacune de leur propre nature.

**Tableau 2 — Catégories des concepts de base de l'évaluation<sup>1</sup>**

Catégorie	Appellation des catégories
Catégorie 1	< Évaluation >
Catégorie 2	< Objet d'évaluation >
Catégorie 3	< Modèle de l'intervention >
Catégorie 4	< Processus d'évaluation >
Catégorie 5	< Critère d'évaluation >

Plus spécifiquement,

- la catégorie < Évaluation > définit le champ de connaissances et de pratique;
- la catégorie < Objet d'évaluation > circonscrit le domaine d'application du champ de connaissances et de pratique;
- la catégorie < Modèle de l'intervention > décrit systématiquement les composantes principales des objets d'évaluation, considérées comme la structure fine du domaine d'application;
- la catégorie < Processus d'évaluation > définit le *modus operandi* de la pratique évaluative;
- la catégorie < Critère d'évaluation > révèle les repères normatifs du processus d'évaluation.

La démarche de conceptualisation nous a permis de délimiter 104 concepts de base de l'évaluation sur lesquels repose le déploiement de la pensée et de la pratique évaluatives.

## b. Identification des termes

Ainsi, après avoir délimité les concepts de base, nous avons attribué un terme (désignation) pour chacun d'eux<sup>2</sup>. Lorsque c'était possible et nécessaire, nous avons opté pour une forme linguistique déjà en usage. Les termes retenus proviennent en bonne partie d'un corpus composé des travaux en évaluation, et particulièrement de la volumineuse production anglo-saxonne de la discipline, traduite ou adaptée en français, et de la documentation francophone, en nette croissance ces dernières années. Nous avons aussi privilégié des mots de la langue générale.

Le tableau 3 présente l'ensemble des termes répartis dans l'une ou l'autre des cinq grandes catégories de concepts de base de l'évaluation.

1. Dans le glossaire, nous avons retenu comme objet d'évaluation l'« intervention » qui convient mieux aux fins du présent manuel consacré aux initiatives de développement. Dans le DME (à paraître), l'objet d'évaluation retenu est « politique », plus englobant, puisque le dictionnaire s'adresse à l'ensemble de la communauté de l'évaluation.

2. Seule exception, le terme *évaluation* réfère à deux concepts distincts : la « discipline » et l'« activité », ce qui explique que nous avons 104 concepts, mais 103 termes.

**Tableau 3 — Répartition des termes dans les cinq catégories de concepts de base de l'évaluation**

Catégorie 1 : < Évaluation >					
Catégorie 2 : < Objet d'évaluation >					
intervention	politique	programme	non-intervention	non-décision	
Catégorie 3 : < Modèle de l'intervention >					
< référentiel >		< mise en œuvre >	< effet >	< valeur de l'intervention >	< acteur, actrice >
raison d'être	objectif	chaîne de production	chaîne d'effets	avantage social	autorité
problème	objectif	production	effet sur cible	coût social	partie prenante
problématique	d'intervention	plan de mise en œuvre	chaîne d'effets sur cible		population cible
modèle de la problématique	chaîne d'objectifs	cible	effet sur cible directe		maître d'œuvre
conception	objectif direct	opérationnelle	effet sur cible intermédiaire		population exposée
intention	objectif intermédiaire	objectif opérationnel	effet sur cible ultime		société
cible	objectif ultime	intransigent	effet hors cible		évaluateur, évaluatrice
cible d'intervention	nature de l'intervention	activité de production	effet intentionnel		expert sectoriel, experte sectorielle
chaîne de cibles	stratégie de l'intervention	extrant	effet non intentionnel		
cible directe	instrument de l'intervention	ressource	effet prévu		
cible intermédiaire	situation hors cible	opération	effet non prévu		
cible ultime	standard	produit résultat			
Catégorie 4 : < Processus d'évaluation >					
< évaluation à priori >	< évaluation in itinéré >	< évaluation à postériori >	< étude >	< produit d'évaluation >	
étude de la raison d'être	étude de la mise en œuvre	étude des effets	observation		
théorie du problème	étude de la conformité	étude de la performance globale	analyse descriptive		
étude de la conception	étude du coût de production		analyse normative		
théorie de l'intervention	coût de production				
	coût unitaire de production				
	étude de la performance opérationnelle				
Catégorie 5 : < Critère d'évaluation >					
< pertinence >	< performance >	< efficacité >	< efficacité >	< équité >	
	performance opérationnelle	efficacité opérationnelle	efficacité opérationnelle	équité opérationnelle	
	performance globale	efficacité globale	efficacité globale	équité globale	

Certaines catégories sont moins développées, tandis que d'autres ont bénéficié d'un niveau supplémentaire de classification. Ainsi,

- les termes de la catégorie < Modèle de l'intervention > sont regroupés dans les sous-catégories < référentiel >, < mise en œuvre >, < effet >, < valeur de l'intervention > et < acteur, actrice >;
- les termes de la catégorie < Processus d'évaluation > sont regroupés dans les sous-catégories < évaluation à priori >, < évaluation in itinéré >, < évaluation à postériori >, < étude > et < produit d'évaluation >;
- les termes de la catégorie < Critère d'évaluation > sont regroupés dans les sous-catégories < pertinence >, < performance >, < efficacité >, < efficience > et < équité >.

### **c. Création de néologismes**

Nous avons créé de nouveaux termes pour répondre à certains besoins conceptuels, comme le permet la démarche terminologique. Ex. : *étude de la raison d'être*, *étude de la conception*, *effet prévu*, *produit d'évaluation*, *évaluation in itinéré*, *non-décision*.

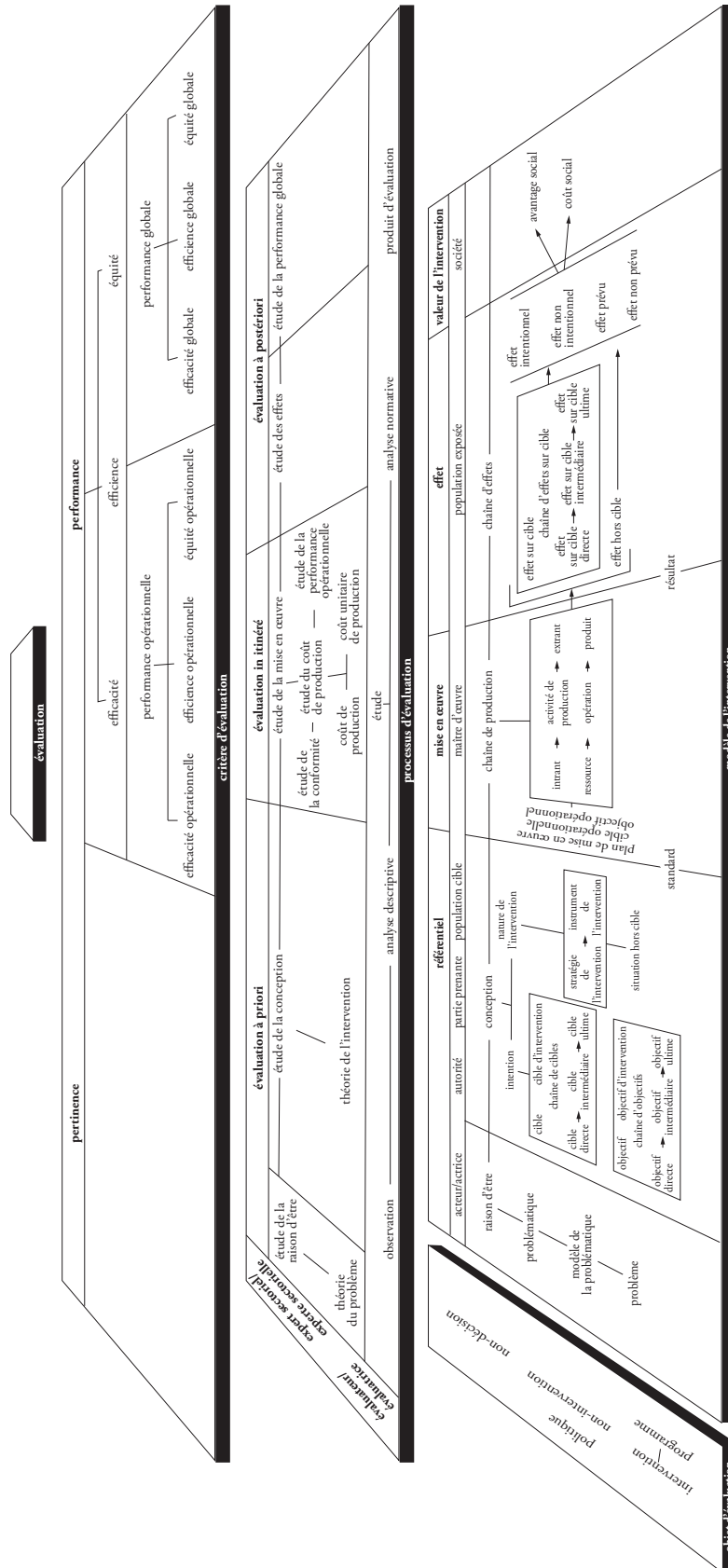
### **d. Classement des concepts dans un système conceptuel**

Le système de concepts, ou système conceptuel, joue un rôle important. Selon l'Organisation internationale de normalisation (2009) :

Un système de concepts sert à clarifier les relations entre concepts, constituer la base d'une terminologie unifiée et normalisée, faciliter l'analyse comparative des concepts et des désignations par-delà les langues et les domaines considérés, faciliter la rédaction des définitions [...].

Dans nos travaux, nous avons fondé le système des concepts de base de l'évaluation sur les relations entre les concepts et, par la suite, nous avons procédé au classement des termes correspondants.

Tableau 4 — Le système des concepts de base de l'évaluation



<https://enap.editionsjfd.com>

Le système de concepts de base de l'évaluation que nous présentons au tableau 4 est un système qui combine plusieurs types de relations. Voici trois exemples de relations entre les concepts :

#### Exemple 1 – Relation hiérarchique générique (espèce-genre)

« cible » : **genre** de plusieurs **espèces** de cible : « cible directe », « cible intermédiaire », « cible ultime ».

#### Exemple 2 – Relation hiérarchique partitive (tout-partie)

« performance » forme un **tout** qui inclut plusieurs **parties** : « efficacité », « efficacité » et « équité ».

#### Exemple 3 – Relation associative

Les objets d'évaluation « politique » et « non-décision » **sont associés entre eux** sans avoir de caractères en commun. Ils marquent des liens thématiques fondés sur l'expérience.

### e. Rédaction des définitions

Enfin, nous avons rédigé des définitions qui permettent de situer chacun des concepts dans une des cinq catégories et de les distinguer des concepts connexes. Toute définition est composée des éléments suivants :

- **le domaine**, qui n'apparaît pas comme tel dans la définition, mais qui participe à l'explication du sens;
- **le définisseur initial**, élément par lequel débute une définition. Il sert à situer le concept à l'étude par rapport aux autres concepts dans un système conceptuel;
- **un ou des caractères essentiels ou distinctifs**, qui sont indispensables à la détermination du concept traité ou qui permettent de distinguer un concept de ceux qui lui sont proches.

Voici deux exemples de définition selon leur composition dans le domaine de l'évaluation :

### Exemple 1 :

Concept : « étude des effets »

Définisseur initial : « étude »

Caractère essentiel : comportant des observations et des analyses descriptives

Caractère distinctif : des effets d'une intervention

Définition complète : Étude comportant des observations et des analyses descriptives des effets d'une intervention.

### Exemple 2 :

Concept : « effet hors cible »

Définisseur initial : « effet »

Caractère essentiel et distinctif : sur une situation hors cible

Définition complète : Effet sur une situation hors cible.

Lorsque le définisseur initial permet d'établir une relation générique avec le concept à définir, il transmet l'ensemble de ses caractères au concept défini. Ainsi, dans l'exemple 1, « étude » transmet l'ensemble de ses caractères à « étude des effets ». La définition de « étude » étant *Examen d'éléments du modèle de l'intervention*, il est superflu de préciser que « étude des effets » est un *examen d'éléments du modèle de l'intervention*, puisque ces caractères sont déjà véhiculés de façon implicite par le définisseur initial « étude ».

## La contribution de la terminologie de l'évaluation à la pratique évaluative

Comme nous l'avons vu précédemment, le système de concepts de base de l'évaluation répartit les cinq grandes catégories de concepts de base de l'évaluation. De chacune de ces catégories découlent des considérations qui encadrent et guident la pratique évaluative, comme nous le verrons dans les lignes qui suivent.

La catégorie 1 < Évaluation > positionne les praticiens de l'évaluation dans un champ disciplinaire unique, distinct des champs voisins.

La catégorie 2 < Objet d'évaluation > circonscrit le domaine où les évaluateurs et évaluatrices peuvent œuvrer. Si le présent ouvrage ainsi que le glossaire en fin de chapitre mettent l'accent sur l'objet d'évaluation intitulé « intervention », le domaine d'application de l'évaluation couvre également les objets d'évaluation « politique », « programme », « non-intervention » et « non-décision ».

La catégorie 3 < Modèle de l'intervention > guide autant le travail du maître d'œuvre que celui des évaluateurs et évaluatrices vers une description détaillée et organisée des initiatives de développement. L'initiative peut être appréhendée par une description de son « référentiel », de sa « mise en œuvre », de ses « effets » et de la « valeur de l'intervention ».

La catégorie 4 < Processus d'évaluation > est dédiée tout particulièrement aux évaluateurs et évaluatrices. Cette catégorie établit les études thématiques qui composent et structurent le processus d'évaluation ainsi que quelques autres concepts incontournables de l'évaluation, telles l'« évaluation à priori », l'« évaluation in itinéré » et l'« évaluation à postériori ».

La catégorie 5 < Critère d'évaluation > contribue à formaliser les questions d'évaluation.

Enfin, l'ensemble des catégories du système des concepts de base contribue à structurer la démarche de type « recherche appliquée » conduisant à des réponses systématiques et rigoureuses aux grandes questions d'évaluation.

D'autres chapitres de cet ouvrage permettent d'approfondir la compréhension des liens entre la terminologie et la pratique. Ainsi, le modèle de l'intervention fait l'objet du chapitre 8 : *La modélisation des interventions en évaluation*. Le chapitre 9 : *Le processus d'évaluation* présente le processus d'évaluation et ses études thématiques. Le chapitre 10 : *Les questions d'évaluation* propose les grandes questions d'évaluation qui découlent des études thématiques du processus d'évaluation. Enfin, les éléments principaux du plan d'évaluation sont développés dans le chapitre 11 : *La matrice d'évaluation*, grâce aux questions d'évaluation.

La terminologie contribue donc à structurer la pratique évaluative et à crédibiliser ses résultats en procédant de l'abstrait (les concepts de base) vers le concret (la matrice d'évaluation), ce qui assure la cohérence et la rigueur de l'évaluation. Ce systématisme dans la pensée évaluative renforce d'autant l'identité propre de l'évaluation.

## **Discussion sur l'apport de différentes langues et cultures à la terminologie de l'évaluation**

Avant de conclure le présent chapitre sur la terminologie de l'évaluation, nous amorcerons une discussion sur la diversité langagière et culturelle dans la construction de la terminologie. Dans un contexte où l'apport de la diversité culturelle dans nos modes de fonctionnement social, politique et économique est valorisé, il y a lieu de se questionner sur l'origine de la terminologie de l'évaluation et sur les contributions de différentes langues et cultures autant dans le passé que dans l'avenir. D'où viennent ces concepts de base en évaluation ? De quelles cultures, de quelles langues, de quelles institutions proviennent-ils ? Quelles sont l'ampleur et les conséquences de ces influences langagières et culturelles ?

À cet égard, nos travaux sur la terminologie nous ont confrontés à certaines observations qui peuvent alimenter la réflexion sur la diversité en évaluation. En voici quelques-unes qui bénéficieraient évidemment d'une validation rigoureuse, mais qui peuvent tout au moins susciter, à titre d'hypothèses, une discussion profitable :

1. L'évaluation des politiques publiques, des interventions ou des programmes publics provient en bonne partie d'un milieu anglo-saxon partageant une langue, une culture et souvent des institutions semblables.



2. La langue de communication internationale dans les instances scientifiques et professionnelles de l'évaluation est principalement l'anglais.
3. Les organisations internationales ont développé une bonne partie de leur corpus conceptuel à partir de la contribution anglo-saxonne à l'évaluation.

Nul ne contestera l'apport de la culture anglo-saxonne à la discipline de l'évaluation et sa grande valeur. L'influence considérable qu'elle a eue sur la discipline se traduit selon toute vraisemblance sur la terminologie et la pensée évaluatives<sup>3</sup>. D'ailleurs, une part importante des 103 termes présentés dans le glossaire de ce chapitre en sont issus.

On peut aisément concevoir comment la langue et la culture anglo-saxonnes ont pu s'imposer en évaluation. La langue anglaise dispose d'un marché intérieur enviable et porte sans conteste le titre de lingua franca dans de nombreux domaines d'échanges internationaux. Les institutions politiques et universitaires, américaines en particulier, ont favorisé l'initiation et le développement de l'évaluation comme champ de connaissances et de pratique. Aujourd'hui, la principale langue d'échange en évaluation au niveau international est la langue anglaise.

La dominance d'une langue et d'une culture, due à un tel contexte, ne peut présumer de la valeur de sa contribution comparativement à celle d'autres langues et cultures. La sensibilité de notre époque à la diversité culturelle s'oppose à l'idée d'une langue ou d'une culture à priori mieux disposée à proposer des concepts pertinents, et nous invite plutôt à porter attention à la richesse présente dans les diverses manifestations langagières et culturelles.

La question centrale devient alors : Quelles sont les conséquences d'une situation où une langue s'impose comme dominante dans une culture scientifique et professionnelle ? La pensée et la langue en évaluation sont-elles alors la résultante de cette langue et de cette culture dominante ? Le cas échéant, cette influence est-elle marginale ou fondamentale ? Et quels sont alors la capacité et l'intérêt des autres langues et cultures à contribuer à la terminologie de l'évaluation ?

Pour répondre à ces difficiles questions, nous utiliserons le cas de la langue et de la culture française, partagée par la francophonie, pour soulever ensuite des enjeux qui concernent plus largement la diversité langagière et culturelle en évaluation.

### *Le cas de la francophonie en évaluation*

La philosophie pragmatique de William James (1911) s'est imposée tôt en évaluation dans le secteur de la *Program Evaluation* et s'est concrétisée par, entre autres choses, le *logic model*, issu de la recherche opérationnelle et par l'accent mis sur les devis expérimentaux et quasi expérimentaux dans la mesure des effets des interventions, au milieu du 20<sup>e</sup> siècle (Marceau, 2016). Ce pragmatisme anglo-saxon s'observe assez simplement dans notre système des concepts de base. Par exemple, les concepts d'« intrant », d'« activité de production », d'« extrant » et d'« effet » concernent la mise en œuvre et les effets des interventions et composent ensemble le *logic model*. Les devis expérimentaux et quasi expérimentaux, qui mettent l'accent sur l'observation plutôt que la déduction, s'appliquent spécifiquement aux effets des interventions.

---

3. L'hypothèse de Sapir-Whorf avance que le langage « forge les dimensions les plus fondamentales de l'expérience humaine ». Le langage et la pensée interagiraient de plusieurs façons significatives (Boroditsky, 2011).

Ces éléments conceptuels, indéniablement pragmatiques, représentent une richesse pour la pensée évaluative. Mais le pragmatisme comporte des limites qu'Émile Durkheim (1914) a mises en évidence peu de temps après la parution des écrits de James :

Comme *méthode*, le Pragmatisme n'est rien d'autre que l'attitude, la tournure générale que doit adopter l'intelligence en présence des problèmes, et cette attitude consiste à tourner nos regards « vers les résultats, les conséquences, les faits ». Pour comprendre le Pragmatisme, dit-il, « [...] il faut seulement le replacer dans son milieu d'origine, le milieu anglo-saxon ».

À cette époque, le pragmatisme est vu comme une menace à la pensée rationaliste, selon le point de vue de Durkheim tout au moins, pour qui la pensée rationaliste est la base même de la culture française. Si la démonstration de Durkheim révèle, en son époque et encore aujourd'hui, des différences marquées dans la recherche de la vérité et des connaissances entre la culture anglo-saxonne pragmatique et la culture française rationaliste, il n'est pas évident d'en conclure que ces philosophies s'opposent plutôt qu'elles ne se complètent dans le contexte d'un champ de connaissances et de pratique.

Nous remarquons d'ailleurs dans notre système de concepts de base la présence de termes que l'on peut associer logiquement à la pensée rationaliste. Les concepts de « référentiel », de « raison d'être », de « problématique », de « modèle de la problématique » et de « théorie de l'intervention », par exemple, se situent en amont des préoccupations de résultats, de conséquences et de faits. Or, ces termes sont peu en usage dans la langue anglaise de l'évaluation. Durkheim a-t-il raison de critiquer la pensée pragmatique pour son insensibilité à la raison, à la théorie, à la préparation à l'action ? James a-t-il raison d'orienter la pensée vers les résultats et les faits ?

Pour notre part, nous avons considéré que ces deux cultures concourent à construire une terminologie et une pensée évaluatives, que ces contributions se complètent et donc que la communauté d'évaluation a avantage à additionner ces contributions plutôt que de les mettre en opposition. Notre système des concepts de base en rend compte par l'intégration ordonnée des concepts dits « pragmatiques » aux concepts dits « rationalistes ».

De cette analyse du cas de la francophonie à la généralisation à la diversité culturelle, il n'y a qu'un pas. Il nous semble en effet fort plausible que la prise en compte de la diversité langagière et culturelle dans l'élaboration des fondements de l'évaluation, et en particulier de sa terminologie, soit grandement profitable à l'évaluation. Pour ce faire, il y a lieu d'être attentifs et de s'interroger sur la valeur des concepts provenant d'autres langues et cultures que celle actuellement dominante en évaluation. Et de s'assurer que l'incorporation de concepts dans la terminologie de l'évaluation se fasse sur la base de leur valeur intrinsèque plutôt que sur la base de leur appartenance à la lingua franca d'aujourd'hui.

## Conclusion

Les travaux terminologiques que nous avons menés ont eu pour ambition première de formaliser les concepts propres à la discipline de l'évaluation. Dans le présent chapitre, nous avons discuté de l'importance de la formalisation en sciences, en général, et en évaluation, en particulier. Nous avons considéré que la terminologie peut contribuer significativement à la formalisation de la discipline de l'évaluation et, conséquemment, à l'affirmation de son identité par rapport aux disciplines connexes. Nous avons élaboré une démarche inspirée de la théorie générale de la terminologie pour former un ensemble de concepts de base de l'évaluation.

L'application de la démarche a consisté à délimiter un ensemble de concepts de base de l'évaluation, à identifier les termes correspondants, à créer des néologismes, à classer les concepts dans un système conceptuel et à rédiger des définitions. Il en est ressorti 103 termes propres au domaine de l'évaluation, un système de concepts de base structuré et des définitions précises qui situent les concepts et les distinguent des concepts connexes.

Nous avons montré que la terminologie de l'évaluation a des incidences concrètes sur la pratique en évaluation. Elle positionne les praticiens dans un champ disciplinaire unique, et circonscrit le domaine d'application de l'évaluation. Elle organise le modèle de l'intervention et ses dimensions principales, c'est-à-dire le référentiel, la mise en œuvre, les effets et la valeur de l'intervention. Elle établit le processus d'évaluation et ses études thématiques desquelles dériveront les questions d'évaluation. Enfin, elle précise les critères d'évaluation.

La terminologie de l'évaluation proposée ici est indéniablement identitaire à l'évaluation. Elle n'est partagée, dans son ensemble, par aucune autre discipline ou pratique professionnelle, aussi connexe soit-elle. Les concepts de base, les termes, le système conceptuel et les définitions qui caractérisent la terminologie de base sont uniques à l'évaluation. La terminologie de l'évaluation proposée ici contribue donc, directement, à l'affirmation de l'identité de la discipline de l'évaluation.

Enfin, il nous apparaît que la diversité langagière et culturelle peut apporter beaucoup à la terminologie de l'évaluation et, en conséquence, à l'évaluation elle-même. Langue et culture sont interdépendantes, et les concepts que chaque langue peut apporter proviennent fort probablement de la culture qui l'accompagne. Cela est une richesse. Ainsi, il semble que les concepts rationalistes que l'on peut associer à la culture francophone contribuent significativement à la terminologie de l'évaluation et complètent fort bien, selon nous, les concepts pragmatiques que l'on peut associer à la culture anglo-saxonne. Plus généralement, la communauté évaluative a intérêt à être attentive aux contributions de diverses langues et cultures et à s'assurer que l'incorporation de nouveaux concepts en évaluation se fasse sur la base de la valeur intrinsèque des idées plutôt que de prendre appui sur la langue principale de communication internationale.

### Messages clés

1. La terminologie, entendue comme l'ensemble des termes propres à un domaine de spécialité, contribue à formaliser les fondements de l'évaluation.
2. Les concepts constituent la base à partir de laquelle les idées prennent vie, se développent et forment un raisonnement. Le terme est l'expression du concept dans un contexte donné.
3. La terminologie favorise le développement de la pensée évaluative.
4. La terminologie contribue à définir et à affirmer l'identité de l'évaluation.
5. La terminologie permet de distinguer le domaine de l'évaluation d'autres domaines de connaissance et de pratique.
6. Les grandes catégories de concepts de base sont l'< évaluation >, l'< objet d'évaluation >, le < modèle de l'intervention >, le < processus d'évaluation > et le < critère d'évaluation >.
7. Chacune des cinq grandes catégories de concepts de base encadre et guide la pratique évaluative.

8. Les concepts de base peuvent être classés dans un système de concepts de base illustrant les relations que les concepts entretiennent entre eux.
9. La définition des termes permet de situer un concept dans une classe d'objets. La définition est formée d'un définisseur initial et d'un ou de plusieurs caractères essentiels ou distinctifs.
10. La diversité langagière et culturelle est profitable à l'évaluation, puisqu'elle peut conduire à incorporer de nouveaux concepts qui ouvrent de nouvelles perspectives.
11. La francophonie a certes beaucoup à apporter à l'évaluation par sa pensée, sa langue et sa culture.

## Exercice de réflexion

L'organisation pour laquelle vous travaillez utilise probablement son propre lexique des termes d'évaluation ou a, tout au moins, ses habitudes langagières. Vous notez certainement des différences dans les termes utilisés par rapport à ceux que nous proposons dans le glossaire qui termine ce chapitre. Partant de là :

1. Est-ce que les différences notées concernent les termes (désignations) ou les concepts (le sens des termes)? Dressez une liste des termes comportant des différences d'appellation et de concept.
2. Les différences notées sont-elles superficielles (désignations différentes) ou plus fondamentales (sens différents)? Indiquez-les.
3. Y a-t-il opposition entre les concepts retenus dans votre organisation et ceux proposés ici? Identifiez les concepts en opposition.
4. À quel point y a-t-il complémentarité dans les deux ensembles de concepts? Faites une liste de tous les termes complémentaires à partir de ceux utilisés dans votre organisation et ceux du glossaire. La liste obtenue contient-elle plus de termes que celle de votre organisation? Si oui, est-ce significatif?
5. Classez les termes de la liste de votre organisation selon les catégories de concepts de base du tableau 3 du présent chapitre. Y a-t-il des catégories mieux représentées grâce à la liste de votre organisation? Concevez-vous d'autres catégories de concepts? Le cas échéant, donnez-leur une appellation.
6. Est-il possible de fonctionner avec plus d'une source de terminologie dans votre milieu de travail? Avez-vous un rôle à jouer dans la terminologie utilisée dans votre organisation?
7. Quels seraient les avantages d'incorporer certaines de nos propositions de termes et de définitions dans votre travail quotidien? Quels seraient les inconvénients?
8. Voyez-vous un avantage à développer d'autres dimensions de la terminologie de l'évaluation que celle des concepts de base, comme la terminologie relative à la méthodologie de l'évaluation, aux techniques de collecte de données, à l'épistémologie de l'évaluation, etc.?

## Pour aller plus loin

**Organisation internationale de normalisation (2009).** *Travail terminologique – Principes et méthodes*, Genève, Suisse, ISO 704:2009.

L'Organisation internationale de normalisation (ISO 704 : 2009) établit et harmonise les principes fondamentaux et les méthodes permettant d'élaborer des terminologies. Elle traite des éléments de base du travail terminologique, tels les concepts, les termes, le système de concepts et les définitions, qui font l'objet du présent chapitre. Nous recommandons la lecture de la norme ISO (704 : 2009), car le présent chapitre en découle.

**Roche, C. (2012).** *Terminologie conceptuelle versus Terminologie textuelle. Repères*, n° 1, janvier.

L'article *Terminologie conceptuelle versus Terminologie textuelle* trace les grandes lignes de la terminologie conceptuelle et de la terminologie textuelle. Dans un texte vulgarisé d'à peine 500 mots, l'auteur résume l'essentiel des deux approches et, du même coup, nous aide à comprendre, de façon sommaire et succincte, la démarche conceptuelle que nous avons suivie dans le présent chapitre. Nous recommandons cet article à tous les lecteurs et lectrices qui désirent un texte qui va droit au but!

**Depecker, L. (2002).** *Entre signe et concept : éléments de terminologie générale*. Presses Sorbonne Nouvelle. <http://books.openedition.org/psn/3395>

Loïc Depecker, linguiste, spécialisé en terminologie et en lexicologie, offre un guide à visée essentiellement pratique et didactique sur la terminologie en général. Il fournit des éléments à la fois théoriques et pratiques, donne des exemples qui permettent une meilleure compréhension des principes amenés et procure une bibliographie dense sur le sujet. Nous le recommandons aux lecteurs et lectrices qui désirent en connaître davantage sur la terminologie en général.

## Références bibliographiques

Boroditsky, L. (2011). How language shapes thought. *Scientific American*, 304(2), 62-65.

Durkheim, É. (1914). *Pragmatisme et sociologie*. Cours inédit prononcé à La Sorbonne en 1913-1914 et restitué par Armand Cuvillier d'après des notes d'étudiants. Collection Les classiques des sciences sociales.

James, W. (1911). *Le pragmatisme*. Introduction par Henri Bergson. Flammarion.

L'Homme, M.-C. (2004). *La terminologie : principes et techniques*. Presses de l'Université de Montréal.

Marceau, R. (2016, décembre). *Mythes et ancrage du pragmatisme en évaluation*. Communication présentée au Carrefour des idées francophones. 2<sup>e</sup> Forum international francophone de l'évaluation. Marrakech, Maroc.

Marceau, R., et Sylvain, F. (2014). *Dictionnaire terminologique de l'évaluation : politiques, programmes, interventions*. Les éditions GID.

Marceau, R., Sylvain, F., Arbour, G. et Hogg, F. (2015, mai). *Réflexion et projet concernant la contribution de la langue française d'évaluation à la pensée évaluative internationale*. Communication présentée lors du Séminaire de recherche francophone de l'évaluation du Réseau Francophone de l'Évaluation. Montréal, Canada.

Marceau, R., Sylvain, F. (à paraître). *Dictionnaire multilingue de l'évaluation (DME)*.

Office québécois de la langue française. (2009). *La rédaction de définitions terminologiques*. Version abrégée et adaptée par Jean Bédard et Xavier Darras.

Organisation internationale de normalisation. (2009). *Travail terminologique – Principes et méthodes*, Genève, Suisse, ISO 704 :2009.

## Glossaire des concepts de base de l'évaluation

Extrait de Marceau, R., Sylvain, F. (à paraître). *Dictionnaire multilingue de l'évaluation (DME)*.

Note : Tout mot souligné dans une définition correspond à un terme défini ailleurs dans le glossaire.

Terme	Définition
<b>acteur, actrice</b>	Personne physique ou morale qui joue un rôle dans une <u>intervention</u> .
<b>activité de production</b>	Activité qui transforme les <u>intrants</u> en <u>extrants</u> .
<b>analyse descriptive</b>	Analyse sur la base d' <u>observations</u> visant à représenter un tout à partir de ses éléments.
<b>analyse normative</b>	Analyse sur la base de <u>critères d'évaluation</u> et de <u>standards</u> visant à porter un jugement.
<b>autorité</b>	<u>Acteur, actrice</u> qui possède le pouvoir et le contrôle de la <u>conception</u> et de la <u>mise en œuvre</u> d'une <u>intervention</u> .
<b>avantage social</b>	<u>Effet</u> positif d'une <u>intervention</u> sur la <u>société</u> .
<b>chaîne d'effets</b>	Série d' <u>effets</u> liés de façon causale.
<b>chaîne d'effets sur cible</b>	Série d' <u>effets sur cible</u> liés de façon causale.
<b>chaîne d'objectifs</b>	Série d' <u>objectifs</u> reliés les uns aux autres qui reflète les <u>intentions</u> de l' <u>autorité</u> .
<b>chaîne de cibles</b>	Série de <u>cibles</u> reliées les unes aux autres qui reflète les <u>intentions</u> de l' <u>autorité</u> .
<b>chaîne de production</b>	Série d'éléments concrets reliés les uns aux autres pour réaliser la <u>mise en œuvre</u> .
<b>cible</b>	Situation devant faire l'objet d'une <u>intervention</u> .
<b>cible d'intervention</b>	<u>Cible</u> relative à la <u>problématique</u> de l' <u>intervention</u> .
<b>cible directe</b>	<u>Cible</u> devant subir en premier lieu l' <u>effet</u> d'une <u>intervention</u> .
<b>cible intermédiaire</b>	<u>Cible</u> devant subir l' <u>effet</u> d'une <u>cible</u> directe.
<b>cible opérationnelle</b>	<u>Cible</u> relative aux <u>opérations</u> de <u>mise en œuvre</u> .
<b>cible ultime</b>	<u>Cible</u> devant subir à terme l' <u>effet</u> d'une <u>intervention</u> .
<b>conception</b>	Processus visant à définir les <u>intentions</u> et la <u>nature</u> de l' <u>intervention</u> .
<b>coût de production</b>	Coût engagé dans la <u>mise en œuvre</u> d'une <u>intervention</u> .
<b>coût unitaire de production</b>	<u>Coût de production</u> divisé par le nombre d' <u>extrants</u> .
<b>coût social</b>	Effet négatif d'une <u>intervention</u> sur la <u>société</u> .
<b>critère d'évaluation</b>	Critère qui sert de base à l' <u>évaluation</u> d'une <u>intervention</u> .
<b>effet</b>	Changement produit par une <u>intervention</u> qui affecte le bien-être de la <u>société</u> .
<b>effet hors cible</b>	<u>Effet</u> sur une <u>situation hors cible</u> .
<b>effet intentionnel</b>	<u>Effet</u> qui correspond aux <u>intentions</u> de l' <u>autorité</u> .
<b>effet non intentionnel</b>	<u>Effet</u> qui ne correspond pas aux <u>intentions</u> de l' <u>autorité</u> .

Terme	Définition
<b>effet non prévu</b>	Effet qui n'est pas anticipé par un <u>acteur</u> , <u>actrice</u> .
<b>effet prévu</b>	Effet qui est anticipé par un <u>acteur</u> , <u>actrice</u> .
<b>effet sur cible</b>	Effet sur une <u>cible</u> .
<b>effet sur cible directe</b>	Effet sur une <u>cible directe</u> .
<b>effet sur cible intermédiaire</b>	Effet sur une <u>cible intermédiaire</u> .
<b>effet sur cible ultime</b>	Effet sur une <u>cible ultime</u> .
<b>efficacité</b>	Performance relative à l'atteinte d' <u>objectifs</u> ou au respect de <u>standards</u> .
<b>efficacité globale</b>	Efficacité de la <u>mise en œuvre</u> et des <u>effets</u> d'une <u>intervention</u> .
<b>efficacité opérationnelle</b>	Efficacité de la <u>mise en œuvre</u> d'une <u>intervention</u> .
<b>efficience</b>	Performance relative à la maximisation de <u>résultats</u> à partir de <u>ressources</u> données.
<b>efficience globale</b>	Efficience de la <u>mise en œuvre</u> et des <u>effets</u> d'une <u>intervention</u> .
<b>efficience opérationnelle</b>	Efficience de la <u>mise en œuvre</u> d'une <u>intervention</u> .
<b>équité</b>	Performance relative à la qualité d'être juste.
<b>équité globale</b>	Équité de la <u>mise en œuvre</u> et des <u>effets</u> d'une <u>intervention</u> .
<b>équité opérationnelle</b>	Équité de la <u>mise en œuvre</u> d'une <u>intervention</u> .
<b>étude</b>	Examen d'éléments du <u>modèle</u> de l' <u>intervention</u> .
<b>étude de la conception</b>	Étude comportant des <u>observations</u> , des <u>analyses descriptives</u> et des <u>analyses normatives</u> de la <u>conception</u> d'une <u>intervention</u> .
<b>étude de la conformité</b>	Étude comportant des <u>observations</u> , des <u>analyses descriptives</u> et des <u>analyses normatives</u> de la conformité de la <u>mise en œuvre</u> d'une <u>intervention</u> avec le <u>plan</u> de mise en œuvre.
<b>étude de la mise en œuvre</b>	Étude comportant des <u>observations</u> , des <u>analyses descriptives</u> et des <u>analyses normatives</u> de la <u>mise en œuvre</u> d'une <u>intervention</u> .
<b>étude de la performance globale</b>	Étude comportant des <u>observations</u> , des <u>analyses descriptives</u> et des <u>analyses normatives</u> de la <u>performance globale</u> d'une <u>intervention</u> .
<b>étude de la performance opérationnelle</b>	Étude comportant des <u>observations</u> , des <u>analyses descriptives</u> et des <u>analyses normatives</u> de la <u>performance opérationnelle</u> .
<b>étude de la raison d'être</b>	Étude comportant des <u>observations</u> , des <u>analyses descriptives</u> et des <u>analyses normatives</u> de la <u>raison d'être</u> d'une <u>intervention</u> .
<b>étude des effets</b>	Étude comportant des <u>observations</u> et des <u>analyses descriptives</u> des <u>effets</u> d'une <u>intervention</u> .
<b>étude du coût de production</b>	Étude comportant des <u>observations</u> et des <u>analyses descriptives</u> du <u>coût de production</u> d'une <u>intervention</u> .
<b>évaluateur, évaluatrice</b>	<u>Acteur</u> , <u>actrice</u> qui possède un haut niveau de compétence en <u>évaluation</u> .
<b>évaluation (sens 1)</b>	Champ de connaissances et de pratique visant à estimer la valeur d'une <u>intervention</u> pour la <u>société</u> .
<b>évaluation (sens 2)</b>	Action d'estimer la valeur d'une <u>intervention</u> pour la <u>société</u> .
<b>évaluation à postériori</b>	Évaluation d'une <u>intervention</u> après sa <u>mise en œuvre</u> .
<b>évaluation à priori</b>	Évaluation d'une <u>intervention</u> avant sa <u>mise en œuvre</u> .
<b>évaluation in itinéré</b>	Évaluation d'une <u>intervention</u> pendant sa <u>mise en œuvre</u> .
<b>expert sectoriel, experte sectorielle</b>	<u>Acteur</u> , <u>actrice</u> qui possède un haut niveau de compétence dans un secteur donné de la <u>problématique</u> de l' <u>intervention</u> .



Terme	Définition
<b>extrait</b>	Résultat de la <u>mise en œuvre</u> destiné à produire les <u>effets</u> .
<b>instrument de l'intervention</b>	Instrument de niveau tactique visant à influencer une <u>cible d'intervention</u> .
<b>intention</b>	Volonté de l' <u>autorité</u> quant à une <u>problématique</u> particulière.
<b>intervention</b>	Politique qui donne lieu à une action sur la <u>société</u> .
<b>intran</b>	Ressource pour l' <u>activité de production</u> .
<b>maître d'œuvre</b>	Acteur, actrice qui est chargé de la <u>mise en œuvre</u> d'une <u>intervention</u> .
<b>mise en œuvre</b>	Processus visant à concrétiser la <u>stratégie de l'intervention</u> .
<b>modèle de l'intervention</b>	Schéma qui présente le <u>référentiel</u> , la <u>mise en œuvre</u> , les <u>effets</u> et la <u>performance globale</u> d'une <u>intervention</u> dans un ordre séquentiel.
<b>modèle de la problématique</b>	Schéma qui présente la <u>problématique</u> dans un ordre causal.
<b>nature de l'intervention</b>	Ensemble des caractéristiques et des propriétés d'une <u>intervention</u> .
<b>non-décision</b>	Politique de ne pas prendre de décision quant à un <u>problème</u> .
<b>non-intervention</b>	Politique qui ne donne pas lieu à une action sur la <u>société</u> .
<b>objectif</b>	<u>Intention</u> d'obtenir un <u>résultat</u> donné sur une <u>cible</u> .
<b>objectif d'intervention</b>	<u>Objectif</u> concernant une <u>cible d'intervention</u> .
<b>objectif direct</b>	<u>Objectif</u> concernant une <u>cible directe</u> .
<b>objectif intermédiaire</b>	<u>Objectif</u> concernant une <u>cible intermédiaire</u> .
<b>objectif opérationnel</b>	<u>Objectif</u> concernant une <u>cible opérationnelle</u> .
<b>objectif ultime</b>	<u>Objectif</u> concernant une <u>cible ultime</u> .
<b>objet d'évaluation</b>	Chacune des <u>politiques</u> d'une <u>autorité</u> .
<b>observation</b>	Action d'examiner et de noter une chose ou un phénomène par des sens, des instruments ou des procédures.
<b>opération</b>	Acte qui transforme des <u>ressources</u> en <u>produits</u> .
<b>partie prenante</b>	Acteur, actrice qui est impliqué dans la <u>conception</u> ou la <u>mise en œuvre</u> d'une <u>intervention</u> ou qui est susceptible d'être touché par les <u>effets</u> d'une <u>intervention</u> .
<b>performance</b>	<u>Critère d'évaluation</u> relatif à des attentes de <u>résultats</u> .
<b>performance globale</b>	<u>Performance</u> de la <u>mise en œuvre</u> et des <u>effets</u> d'une <u>intervention</u> .
<b>performance opérationnelle</b>	<u>Performance</u> de la <u>mise en œuvre</u> d'une <u>intervention</u> .
<b>pertinence</b>	<u>Critère d'évaluation</u> relatif à la qualité d'être approprié à une situation donnée.
<b>plan de mise en œuvre</b>	Plan comportant des processus et des <u>opérations</u> pour la <u>mise en œuvre</u> d'une <u>intervention</u> .
<b>politique</b>	Ligne de conduite d'une <u>autorité</u> quant à un <u>problème</u> .
<b>population cible</b>	Population qui doit être exposée à une <u>intervention</u> .
<b>population exposée</b>	Population qui a été exposée à une <u>intervention</u> .
<b>problématique</b>	Ensemble des caractéristiques, des causes et des conséquences d'un <u>problème</u> .
<b>problème</b>	Situation jugée insatisfaisante pour le bien-être d'une population.
<b>processus d'évaluation</b>	Processus visant à concrétiser une <u>évaluation</u> par la conduite d'une série d' <u>études</u> .
<b>produit</b>	Résultat d'une <u>opération</u> de la <u>mise en œuvre</u> .
<b>produit d'évaluation</b>	Résultat d'un <u>processus d'évaluation</u> .
<b>programme</b>	<u>Intervention</u> prenant la forme d'une série d'activités concrètes.



Terme	Définition
<b>raison d'être</b>	Ensemble des motifs qui légitime une <u>intervention</u> .
<b>référentiel</b>	Ensemble d'éléments de la <u>raison d'être</u> et de la <u>conception</u> qui sert à orienter la <u>mise en œuvre</u> d'une <u>intervention</u> .
<b>ressource</b>	Support pour une <u>opération</u> .
<b>résultat</b>	Ce qui est produit dans le cadre d'une <u>intervention</u> .
<b>situation hors cible</b>	Situation qui ne figure pas parmi les <u>cibles de l'intervention</u> .
<b>société</b>	Population d'une région ou d'un pays donnés qui partage des lois, des institutions et des <u>politiques</u> .
<b>standard</b>	Valeur d'un <u>critère d'évaluation</u> basé sur une référence exemplaire.
<b>stratégie de l'intervention</b>	Stratégie consistant à coordonner les <u>instruments de l'intervention</u> pour atteindre les <u>objectifs d'intervention</u> .
<b>théorie de l'intervention</b>	Construction de l'esprit visant à décrire l' <u>intervention</u> et à dégager des principes normatifs.
<b>théorie du problème</b>	Construction de l'esprit visant à décrire le <u>problème</u> et à dégager des principes normatifs.
<b>valeur de l'intervention</b>	Valeur d'une <u>intervention</u> pour la <u>société</u> .



# Les objets d'évaluation en contexte de développement international



*Jean Serge Quesnel*  
*Marie-Hélène Adrien*

## Introduction

La première question qui est posée au moment du lancement d'une évaluation est de connaître l'objet sur lequel va porter la démarche. L'objet signifie ce que l'on entreprend d'évaluer. La seconde question porte sur la faisabilité d'évaluer l'objet ciblé, et la troisième question est de savoir comment on devra procéder pour réaliser l'évaluation. Puis, il faut s'intéresser à la communication et à l'utilisation des résultats de l'évaluation. Dans ce chapitre, nous nous attarderons à considérer la première question. Les autres chapitres de l'ouvrage répondent aux autres questions. L'évaluation requiert la construction d'un jugement découlant d'une démarche rationnelle. Elle suppose donc une transformation des représentations permettant d'interpréter des éléments factuels sur la base d'une grille de critères (Barbier, 1985). La construction du jugement est au cœur de la connaissance évaluative. Évidemment, ceci débute par un bon discernement de l'objet d'évaluation. Tout peut faire l'objet d'une évaluation.

## Objets d'évaluation

L'univers des objets d'évaluation est vaste et complexe. On peut aussi bien évaluer un concept, une action, un produit, un service, une intervention, un organisme, que des situations, des politiques, des tactiques, des stratégies, des projets, des programmes, des pratiques et des processus. On peut, de même, évaluer des phénomènes, des comportements, des valeurs morales et économiques, des pratiques sociales, des théories, des modèles ou des plans. Dans le cadre d'une évaluation du changement, l'état initial des lieux, les diverses transformations résultant de la progression de l'implantation de l'initiative de développement, les résultats finaux et leurs retombées constituent autant d'objets d'évaluation (Brousselle *et al.*, 2011).

Dans le domaine du développement international, l'objet d'évaluation est le plus souvent une intervention. Celle-ci peut être conçue comme un système organisé d'actions qui, dans un environnement donné, comprend : des agents (les acteurs), un contexte qui échappe au contrôle des acteurs, une structure (l'ensemble des ressources et des règles), des processus (relations entre les ressources et les activités) et un but (l'état futur vers lequel le processus d'action est orienté). Les résultats de l'intervention sont le produit de longues chaînes causales qui mettent en relation l'ensemble des structures, les processus et les retombées (Parsons, 1977). Dans la pratique du développement international, les

objets d'évaluation des interventions de développement englobent les actions, les projets, les programmes, les stratégies et les politiques de développement international, chacun de ces objets constituant ainsi une unité d'analyse.

### **Les intentions comme objets d'évaluation**

L'objectif de l'évaluation d'une intervention de développement varie en fonction de l'intention, comme illustré ci-dessous :

- Évaluer l'évaluabilité de l'intervention (Beaudry et Gauthier, 1992). Il s'agit de l'étape préliminaire permettant de vérifier qu'il est possible de réaliser l'évaluation;
- Évaluer l'atteinte des résultats espérés;
- Évaluer la satisfaction des parties prenantes;
- Évaluer les effets environnementaux;
- Évaluer la pertinence du devis d'intervention;
- Évaluer la validité de la théorie de changement;
- Évaluer la cohérence de l'intervention et sa connectivité avec d'autres interventions;
- Évaluer la pérennité des retombées;
- Évaluer l'équité de l'intervention;
- Évaluer l'efficacité des modalités d'intervention;
- Évaluer l'efficience des moyens pris et l'utilisation des ressources;
- Évaluer le coût-bénéfice d'une intervention;
- Évaluer la contribution des intervenants;
- Évaluer une intervention à des fins de reddition de comptes et d'apprentissage.

### **Les objets d'évaluation en fonction des modes d'évaluation**

Les évaluations d'intervention de développement peuvent se décliner sur plusieurs modes :

- En mode autoévaluation (évaluer par soi-même ses propres résultats);
- En mode participatif (évaluation qui engage les parties prenantes);
- En mode indépendant (évaluation conduite par des services non liés aux responsables);
- En mode conjoint (évaluation à laquelle collaborent différents organismes);
- En mode multipartite (évaluation à laquelle contribuent plusieurs acteurs);
- En mode métaévaluation (agrégation des trouvailles d'une série d'évaluations);
- En mode évolutif (évaluation d'un processus évolutif dans un contexte de changement).

## Objets d'évaluation du développement international

Les interventions de développement international sont interreliées et multidimensionnelles. Elles comprennent des initiatives simples, complexes, voire extrêmement complexes qui obligent, à chacun des niveaux présentés ci-dessous, une application de concepts singuliers, lesquels présentent des défis méthodologiques. Pour structurer l'approche évaluative des interventions de développement international, nous distinguons six niveaux d'objets d'évaluation :

Premier niveau : Les actions menant aux transferts d'intrants de développement telles que la fourniture de denrées, de biens et de services, la réalisation d'une campagne d'immunisation et le développement des capacités par l'envoi d'enseignants expatriés.

Deuxième niveau : Les extraits de développement ciblés que sont les projets. Par exemple : la construction d'une infrastructure ou une intervention de mesures d'urgence.

Troisième niveau : Les interventions réalisées en partenariat produisant des retombées développementales. Par exemple : le développement d'un secteur forestier, agricole, minier, ou encore un programme de coopération à l'échelle d'un pays.

Quatrième niveau : L'efficacité des acteurs organisationnels du développement. Ce niveau comprend des évaluations tels le renforcement des capacités institutionnelles ou des initiatives de développement multipartites à l'échelle nationale, régionale et mondiale.

Cinquième niveau : Les stratégies d'intervention misant sur des impacts de développement, telles que les stratégies d'allègement de la pauvreté, les ajustements structurels, les cibles du développement durable et les stratégies d'amélioration de l'efficacité de l'aide au développement.

Sixième niveau : Les politiques permettant l'optimisation des interventions développementales, telles que les initiatives visant la mise en œuvre de politiques, par exemple une politique sur la jeunesse ou encore les politiques d'équité, les plateformes politiques et économiques et les politiques de développement.

## Évolution historique des niveaux d'objets d'évaluation du développement international

Chaque niveau d'objets d'évaluation reflète une période de l'évolution de la gestion du développement international. À ses débuts, dans les années 1950, la gestion du développement consistait à exporter des denrées, des biens et des services de coopération technique, ce qui exigeait une gestion axée sur le transfert d'*intrants de développement*, correspondant au premier niveau d'objets d'évaluation. Au début des années 1960, alors que l'aide au développement international était liée à la fourniture de biens et services achetés dans le pays donateur, un examen approfondi a révélé un gaspillage important des intrants expédiés internationalement. À l'époque, les capacités d'absorption des pays bénéficiaires n'étaient pas prises en compte. Par exemple, une flotte de véhicules était envoyée sans que les ressources nécessaires à l'entretien des équipements n'aient été prévues.

Au cours des années 1970, afin d'assurer une meilleure reddition de comptes et une plus grande efficacité et efficience de l'aide apportée, la communauté du développement international a adopté la gestion par projet, ce qui correspond au deuxième niveau d'objets d'évaluation. Axé sur l'atteinte

d'*extrants de développement*, ce type de gestion mise sur un contrôle plus serré des paramètres d'intervention. Idéalement, toutes les composantes nécessaires à la réussite d'une intervention relèvent du gestionnaire de projet, et ce, dans le but de minimiser les risques (économiques, sociaux, politiques ou environnementaux).

Au cours de cette période, une surabondance de financement a permis de mettre en œuvre une pléthore d'interventions de développement fondées sur la gestion par projet. Les résultats ne se sont pas fait attendre : la pérennité et l'appropriation nationale des interventions réalisées ont été rapidement compromises par des dysfonctionnements de la pratique de l'aide au développement et par une vision étroite de résultats à court terme. Dans une critique intitulée *Trouver des problèmes aux solutions*, le Club du Sahel<sup>1</sup> a exprimé le besoin d'un changement de paradigme. Les pays du Sahel éprouvaient un sentiment d'insatisfaction et de lassitude à l'égard de l'aide reçue. Dans la course à l'efficacité, il n'y avait pas de vainqueur mais un perdant : le pays bénéficiaire (Naudet, 1999).

La coopération internationale était en crise. Son efficacité, voire son utilité, était remise en question. Des réformes étaient nécessaires pour inviter les bailleurs à collaborer et à être moins visibles afin d'avoir, ensemble, un plus grand impact durable de développement. L'aide devait retrouver sa juste place au cœur des besoins et au service du développement des pays concernés. Au même moment, des crises mondiales se sont succédé, dont celle de l'endettement. Les institutions financières internationales ont alors lancé l'aide-programme, accompagnée de mesures d'ajustement structurel, et les bailleurs de fonds se sont serré les coudes pour collaborer au financement et à l'opérationnalisation de programmes de redressement socio-économiques, amorçant ainsi l'avènement du troisième niveau d'objets d'évaluation, la collaboration multipartite.

Malgré les milliards de dollars investis au fil des années, force a été de constater que les progrès au niveau national ne suivaient pas. La conclusion s'imposait d'elle-même, le développement d'un pays ne peut pas être propulsé de l'extérieur, mais doit émaner de l'intérieur. Il doit donc se fonder sur une bonne gouvernance des institutions. Dans cet esprit, la *Déclaration de Paris sur l'efficacité de l'aide* (OCDE, 2005) recentre le développement à partir des pays, sur la notion de partenariat. Conçus dans un esprit de responsabilité mutuelle, les engagements de partenariat reposent sur les principes d'appropriation, d'alignement, d'harmonisation et de gestion axée sur les résultats. Bien que l'évaluation institutionnelle, le quatrième niveau d'objets d'évaluation, existe depuis les années 1990, la Déclaration a provoqué une prise de conscience sur le besoin de développer les capacités institutionnelles des pays.

Le cinquième niveau d'objets d'évaluation est celui des stratégies d'intervention de développement. Il s'est cristallisé autour des *Objectifs du millénaire pour le développement* (ONU, 2000), suivis des *Objectifs du développement durable* (ONU, 2015). Pour la première fois dans l'histoire du monde, en l'an 2000, les 193 pays membres des Nations Unies se sont engagés à atteindre des cibles destinées à améliorer le bien-être de l'humanité à l'échelle de la planète. Quinze ans plus tard, ces mêmes pays se sont engagés à redoubler d'efforts pour atteindre les cibles de développement durable. Les objectifs de développement durable donnent la marche à suivre pour offrir un avenir meilleur et plus durable à tous. Ils répondent aux défis mondiaux, notamment ceux liés à la pauvreté, aux inégalités, au climat,

---

1. Le Club du Sahel et de l'Afrique de l'Ouest (CSAO) est une plateforme internationale indépendante. Son Secrétariat est hébergé au sein de l'Organisation de coopération et de développement économiques (OCDE). Sa mission est de promouvoir des politiques régionales à même d'améliorer le bien-être économique et social des populations du Sahel et de l'Afrique de l'Ouest. ([www.oecd.org/csao](http://www.oecd.org/csao))

à la dégradation de l'environnement, à la prospérité, à la paix et à la justice. Les objectifs sont interconnectés, et il est important de réaliser chacun d'entre eux et d'atteindre chacune des cibles qui misent sur des retombées (*impacts*), d'ici 2030.

Enfin, le sixième niveau d'objets d'évaluation est celui des politiques. Il s'agit d'évaluer des politiques publiques articulées aux niveaux local, régional, national et global, une pratique que l'ONU, l'OCDE (aide au développement) et l'Union européenne promeuvent depuis plus de 20 ans sur le plan international. L'évaluation d'une politique publique a pour objet d'apprécier l'efficacité des politiques de développement en comparant les résultats aux objectifs et aux moyens mis en œuvre.

Les finalités de l'évaluation sont les suivantes : améliorer l'efficacité des services, mieux prendre en compte les attentes et l'intérêt des populations, améliorer l'adéquation de la politique à son contexte, quantifier les résultats, améliorer l'allocation des ressources, responsabiliser les acteurs et éviter l'institutionnalisation de mesures inadéquates. Les critères utilisés pour l'évaluation d'une politique sont : la pertinence (problèmes/objectifs), la cohérence (problèmes/objectifs), l'efficacité (objectifs/effets), l'efficience (effets/moyens), l'impact systémique (extrants/réalisations) et l'utilité. Une évaluation de politique publique (EPP) repose sur une étude exhaustive de la documentation disponible et des travaux déjà produits, et sur des techniques d'analyse (statistiques, économétrie, sciences des données, coûts, impacts, effets de levier...), d'enquêtes (entretiens, déplacements de terrain) et de consultation (questionnaires, sondages, groupes de discussion). Elle vise *l'optimisation* des interventions de développement.

## Modalités utilisées à chaque niveau d'objets d'évaluation

### Premier niveau : objets d'évaluation des résultats-intrants de développement

Le niveau d'objets d'évaluation des résultats-intrants de développement correspond aux évaluations des tâches administratives et bureaucratiques. L'objectif est l'accomplissement de la tâche, dans le respect des règles. Autrement dit, la procédure, soit l'approche systémique, détermine la qualité du livrable. La façon la plus rapide d'exécuter une tâche de qualité est donc de faire en sorte que toutes les étapes administratives soient effectuées dans le respect de la procédure en place, ce qui ne laisse que peu de place à une certaine autonomie de gestion.

L'intention d'évaluation à ce premier niveau est ainsi résumée *a contrario* par Osborne et Gaebler (1992) :

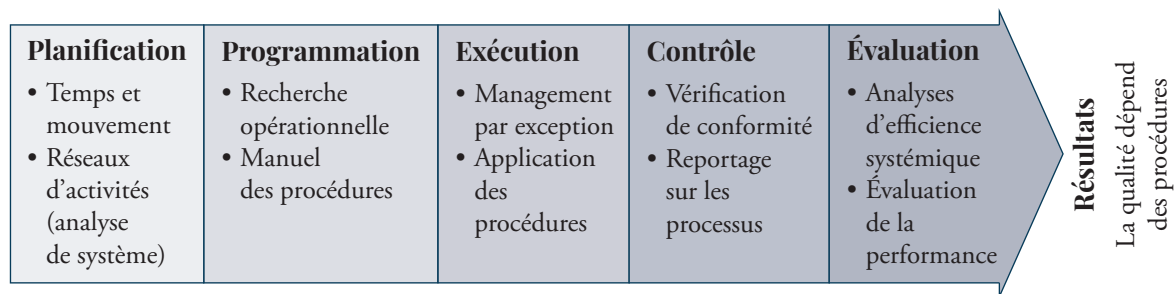
- Si vous ne mesurez pas les résultats, vous ne pouvez pas distinguer le succès de l'échec;
- Si vous ne pouvez pas déceler le succès, vous ne pouvez pas le récompenser;
- Si vous ne récompensez pas le succès, vous récompensez probablement l'échec;
- Si vous ne voyez pas le succès, vous ne pouvez pas apprendre de lui;
- Si vous ne pouvez pas déceler l'échec, vous ne pouvez pas le corriger.

Ce premier niveau repose sur une théorie d'organisation du travail selon laquelle le flux des tâches est analysé et synthétisé dans le but principal d'améliorer l'efficience des processus des produits livrables. Cette théorie, développée par Frederick Winslow Taylor, a fait son apparition au cours des

années 1880 et 1890 au sein des industries manufacturières, et a connu son apogée vers 1910. Elle a été l'une des premières tentatives d'application de principes scientifiques à l'ingénierie, aux processus et à la gestion.

À ce niveau, l'objet d'évaluation peut être associé à la surveillance et à l'amélioration des procédures. Aujourd'hui, cette organisation du travail est désignée sous le terme de « taylorisme digital ». Il s'agit de maximiser l'efficacité et l'efficacités par la standardisation des outils et techniques nécessaires pour accomplir chaque tâche afin de livrer un produit de qualité. Chaque tâche est analysée pour identifier les mouvements essentiels, puis est transposée dans une procédure exacte. Cette vision mécaniste met l'accent sur la prévisibilité et l'uniformité tout en réduisant les erreurs. Il est donc plus facile pour les gestionnaires de maintenir le niveau de productivité en exerçant des pressions sur les employés. Il suffit de penser à Amazon, à Costco ou à McDonald's.

Le schéma qui suit résume les divers éléments du cycle de gestion des résultats-intrants de développement :



**Figure 1**

Niveau de gestion et d'analyse évaluative des résultats intrants au développement.

À ce premier niveau, les efforts d'évaluation portent normalement sur les considérations d'efficacité, d'efficacité, de qualité du livrable et sur la satisfaction du client.

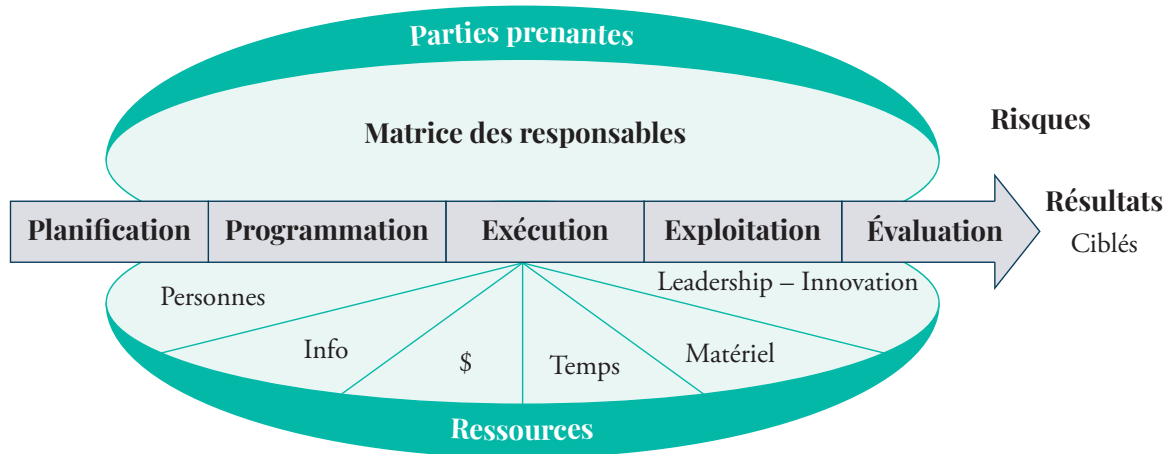
## Deuxième niveau : objets d'évaluation des résultats-extrants de développement

Le niveau d'objets d'évaluation des résultats-extrants de développement est celui des projets. À ce niveau, l'objectif est d'atteindre un résultat ciblé. Cette approche cherche à maximiser les chances de succès par le contrôle du plus grand nombre de paramètres possible afin de ne laisser que peu de place à l'incertitude et ainsi réduire les risques d'échec. Une multitude de livres ont été publiés sur l'art et la science de gérer des projets, mais la principale référence demeure le *PMBOK Guide* (PMI, 2021).

Selon le *Project Management Institute* (PMI), un projet est un effort temporaire consenti dans le but de créer un résultat unique. Si le caractère temporaire d'un projet implique un commencement et une fin, cela ne signifie pas qu'il est de courte durée. La fin est atteinte lorsque les objectifs du projet sont réalisés ou lorsque le projet est arrêté parce que ses objectifs ne seront pas atteints ou ne peuvent pas l'être, ou encore parce que le projet n'est plus utile. Par ailleurs, il convient de distinguer le projet de son résultat. Si la nature du premier est temporaire, celle du second se veut durable. D'ailleurs, la durée de l'impact social, économique et environnemental peut être plus longue que les projets eux-mêmes.



Le schéma qui suit résume les facteurs qui entrent en ligne de compte dans la réalisation d'un projet.



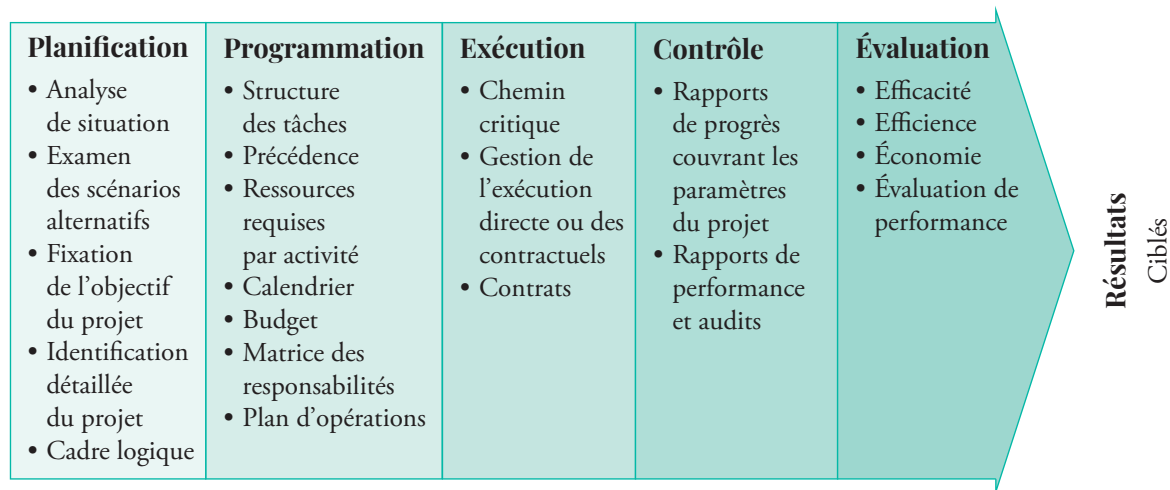
**Figure 2**

Univers de gestion et d'analyse évaluative de l'intervention, projet de développement.

La gestion de projet consiste à gérer les ressources (humaines, financières, technologiques et de temps) de façon ordonnée et économique en vue de réaliser un objectif prédéfini, en conformité avec les paramètres relatifs aux spécifications techniques, aux coûts et aux délais, et ce, dans un contexte où les risques abondent.

L'interrelation entre ces facteurs est telle qu'un changement dans l'un d'eux risque d'affecter au moins un des autres facteurs. Par exemple, un raccourcissement des délais entraîne souvent une augmentation du budget afin d'obtenir les ressources supplémentaires nécessaires à l'exécution du même travail en moins de temps. S'il n'est pas envisageable de dégager des fonds, le contenu ou la qualité du projet s'en ressentira. Cette interrelation joue également sur des ramifications potentiellement plus délicates à gérer, lorsque les parties intéressées d'un projet ont des idées différentes sur l'importance relative des facteurs. La modification des paramètres peut engendrer des risques supplémentaires. L'équipe de projet doit donc être capable d'évaluer la situation et d'équilibrer les demandes.

Le schéma qui suit résume les divers éléments du cycle de gestion des résultats-extrants du développement :



**Figure 3**

Niveau de gestion et d'analyse évaluative des résultats – extrants pour le développement.

À ce deuxième niveau, les efforts d'évaluation portent sur les considérations d'économie, d'efficience, d'efficacité, de pertinence, de qualité et de pérennité des résultats atteints et leurs effets.

### Troisième niveau : objets d'évaluation des résultats-retombées de développement

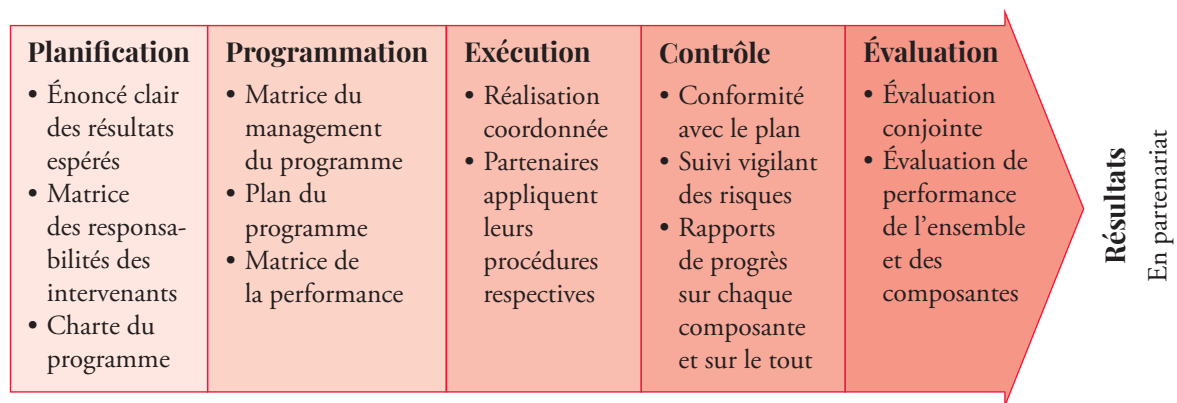
Le niveau d'objets d'évaluation des résultats-retombées de développement est celui des interventions-programmes d'envergure réalisées en partenariat. Un programme est un ensemble d'actions visant l'accomplissement d'une intention claire et comprenant des indications sur ce qui doit être fait, par qui, de quelle façon et avec quelles ressources. Il s'agit d'un effort conjoint destiné à produire d'importantes retombées avantageuses pour le pays bénéficiaire.

La gestion par programme est plus appropriée lorsque l'ampleur de l'intervention, l'effort et les ressources requises excèdent la capacité d'une seule entité. Elle se caractérise par un niveau élevé d'interrelations et d'interdépendance entre les intervenants de différentes entités et les acteurs sociétaux. Elle représente un enjeu stratégique majeur qui touche directement à la mission et à l'engagement des entités concernées.

Si la gestion de projet est un système fermé, résumé dans son cadre logique (Sansom, 2015), les interventions en partenariat sont, elles, des systèmes ouverts pluripartites (multilatéraux), illustrés par des modèles logiques ouverts et multidimensionnels.

Un partenariat peut se définir comme une association de différents intervenants. Chacun conserve son autonomie, mais ils acceptent de joindre leurs efforts en vue de réaliser un objectif commun relié à un besoin clairement identifié envers lequel, en vertu de leurs missions respectives, ils ont un intérêt, une motivation, une responsabilité (Frank et Smith, 2020).

Le schéma qui suit résume les divers éléments du cycle de gestion des résultats-retombées de développement.



**Figure 4**

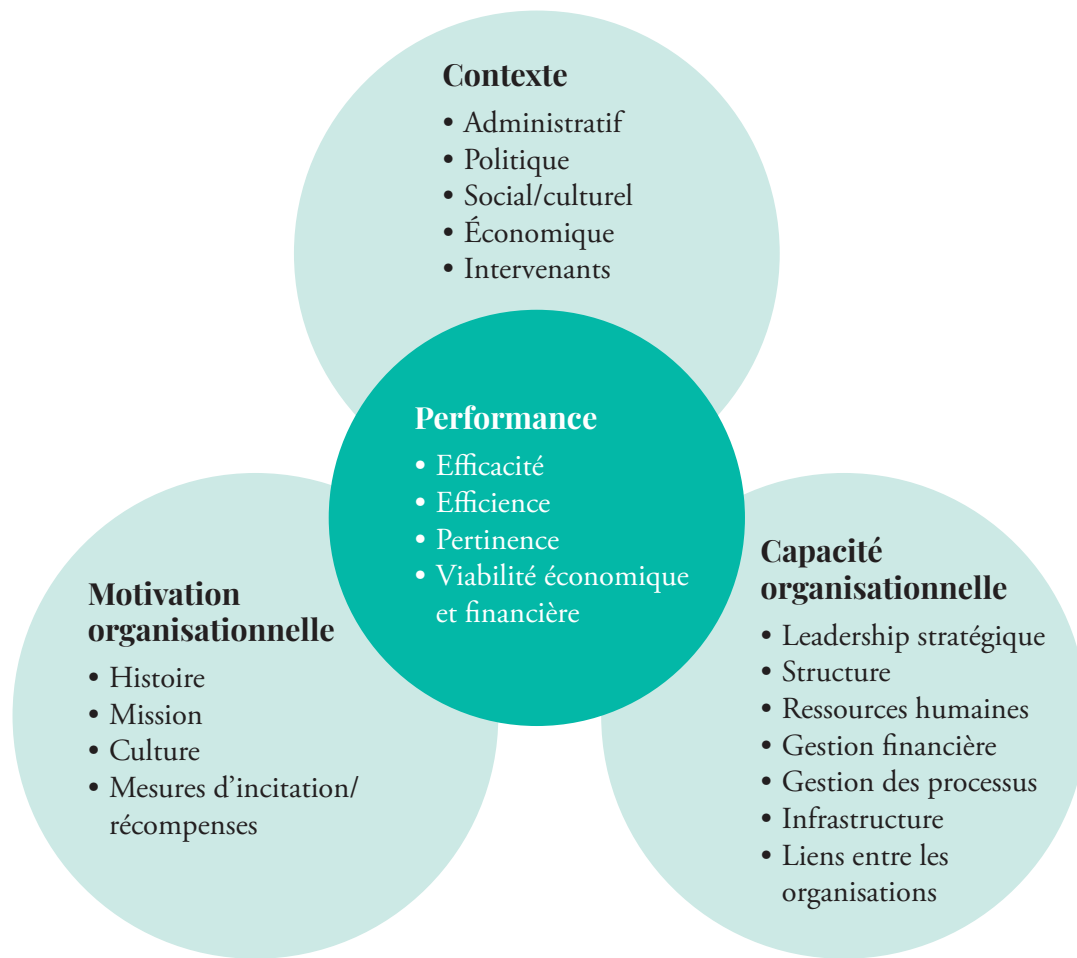
Niveau de gestion et d'analyse évaluative des résultats-retombées de développement.

Les efforts d'évaluation à ce troisième niveau d'objets d'évaluation portent sur les considérations d'efficacité, de pertinence, de connectivité, de qualité des résultats-retombées atteints, d'appropriation et d'habilitation.

#### **Quatrième niveau : objets d'évaluation de l'efficacité des acteurs livrant les résultats de développement**

Au cours des années 2000, la communauté internationale a compris que les principaux facteurs de développement sont la capacité institutionnelle d'un pays à s'autogérer et la qualité des acteurs institutionnels qui appuient les efforts de développement. Dans ce contexte, il est impératif d'être en mesure d'évaluer les capacités institutionnelles qui contribuent au développement et le stimulent.

Le Centre de recherches pour le développement international (CRDI) et la Banque interaméricaine de développement (BID), avec l'appui d'Universalialia (Lusthaus *et al.*, 2003), proposent un modèle simple et complet pour poser un diagnostic institutionnel. Le schéma ci-dessous reprend les variables à examiner pour établir un diagnostic organisationnel.



**Figure 5**

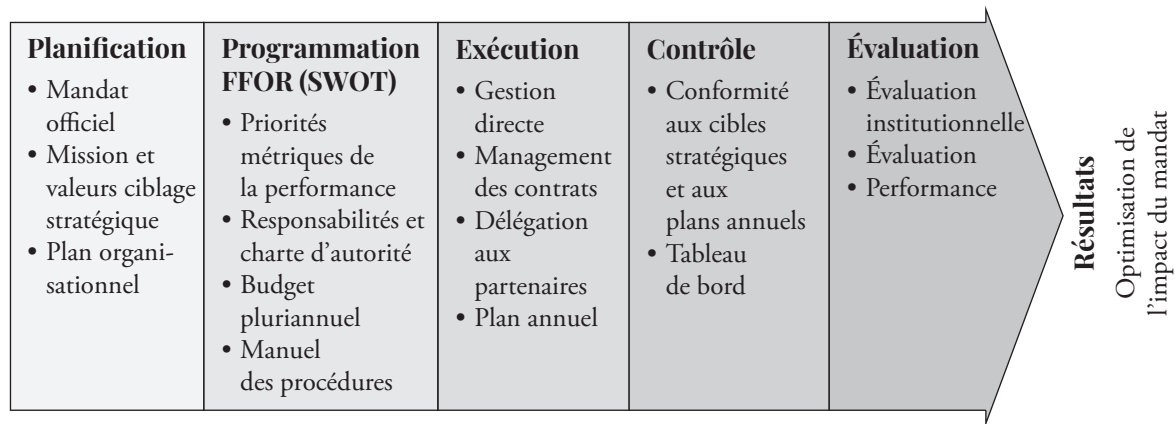
Niveau de gestion et d'analyse évaluative de l'efficacité des acteurs livrant les résultats de développement.

Source : Adapté de Lusthaus *et al.*, 2003, p. 36.

Cet extrait a été reproduit aux termes d'une licence accordée par COPIBEC.

Ce modèle est fréquemment utilisé par plusieurs organisations de développement.

Le schéma qui suit indique les outils de bonne gestion utiles à l'analyse évaluative d'une organisation.



**Figure 6**

Niveau de gestion et d'analyse évaluative des acteurs livrant les résultats de développement.

Les efforts d'évaluation à ce quatrième niveau portent sur les considérations de la performance organisationnelle (efficacité, pertinence, viabilité financière), des capacités organisationnelles, de la motivation organisationnelle et du contexte habilitant.

De telles évaluations sont réalisées de façon conjointe afin de réduire le fardeau des activités d'évaluation et d'éliminer la duplication; de donner l'occasion d'échanger et d'apprendre; d'encourager une diversité de perspectives et de talents. Les évaluations institutionnelles multipartites permettent de réaliser des mandats de plus grande envergure dotés de ressources accrues mises en commun. Cela produit des évaluations de meilleure qualité, plus crédibles, dont les constats sont mieux reçus et qui exercent une plus grande influence sur la prise de décision des parties prenantes.

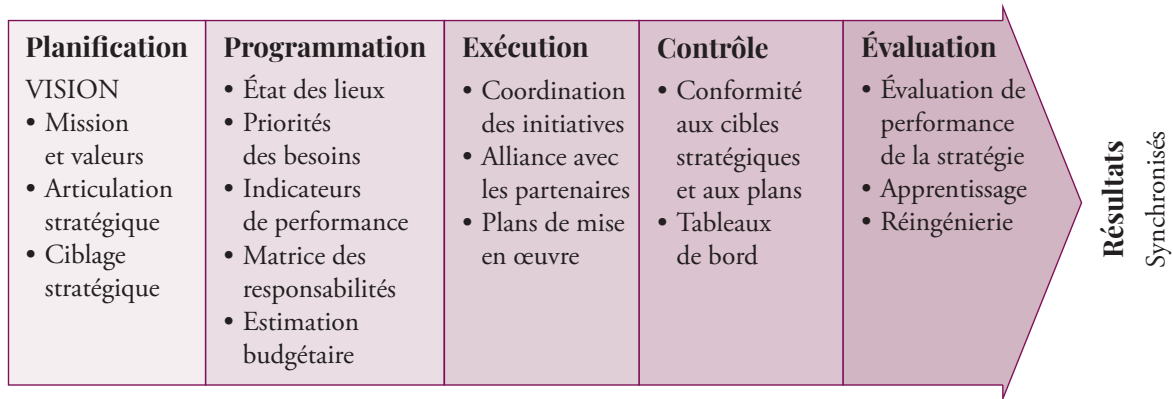
### Cinquième niveau : objets d'évaluation des résultats-impacts de développement

Le cinquième niveau des résultats-impacts de développement est celui des stratégies d'intervention de développement. Une stratégie est l'art de combiner et de coordonner diverses actions pour atteindre un but. Une stratégie axée sur les résultats de développement doit pouvoir : énoncer de façon réaliste les cibles; identifier et engager les parties prenantes pour répondre à leurs besoins; assurer un suivi à l'aide d'indicateurs *SMART*<sup>2</sup> de rendement; identifier et gérer les risques; enseigner par l'action; intégrer l'apprentissage dans les décisions de poursuite.

Les exemples de stratégies mondiales les plus connues de la communauté du développement international sont celles adoptées par les 192 pays membres des Nations Unies, soient les objectifs du Millénaire pour le développement (OMD) en 2000 et les objectifs de développement durable en 2015.

2. SMART – Sigle anglophone qui signifie que les indicateurs sont simples, mesurables, atteignables, réalistes et limités dans le temps.

Le schéma ci-après illustre le processus d'élaboration, de programmation, d'implantation, de suivi et d'évaluation.



**Figure 7**

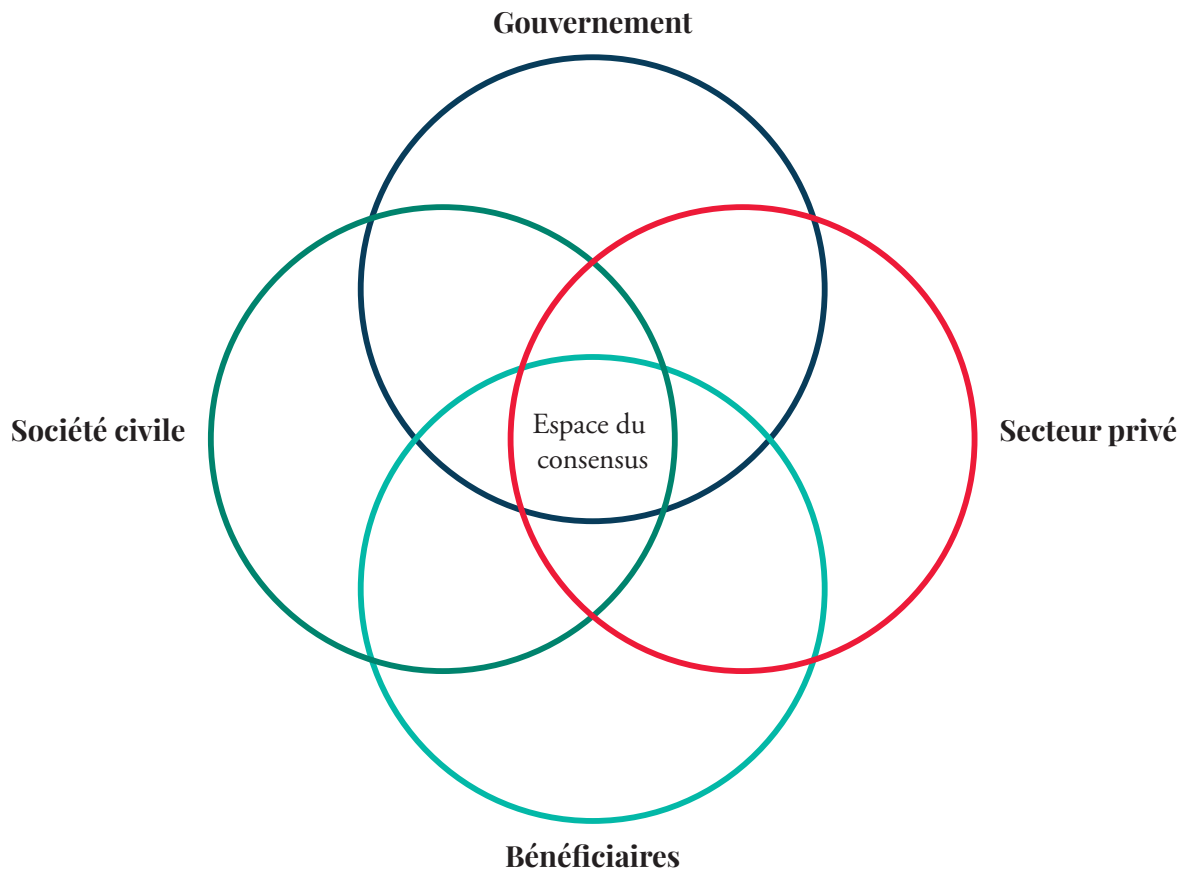
Niveau de gestion et d'analyse évaluative des résultats-impacts de développement.

À ce cinquième niveau des résultats-impacts de développement, les efforts d'évaluation portent sur l'atteinte des impacts espérés, non prévus et non souhaitables, les progrès accomplis au regard des indicateurs clés et l'explication des écarts de rendement sur l'équité des impacts.

### Sixième niveau : objets de l'évaluation de l'optimisation du développement

Le sixième niveau des résultats-optimisation de développement est celui des politiques de développement. Une politique de développement est l'aboutissement d'un processus décisionnel interactif, délibéré, motivé par le désir d'atteindre des résultats avantageux pour le bien commun. Il s'agit d'une construction sociale dont la structure et la finalité intègrent les intérêts communs et qui scellent la collaboration entre les acteurs socio-économiques.

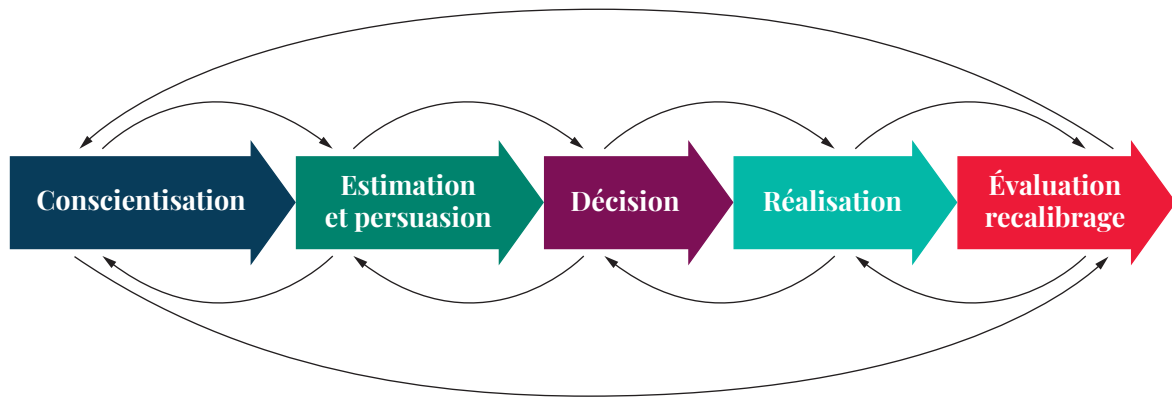
Pour qu'une politique soit l'expression d'un consensus entre les parties prenantes, elle doit laisser une certaine marge de manœuvre à ces dernières pour l'adoption et l'application de la politique en question, comme le montre le schéma ci-dessous. Si une politique ne promeut que les intérêts de l'une des parties prenantes, sa mise en œuvre risque non seulement de mal se dérouler, mais elle pourrait rencontrer de la résistance, voire de l'opposition.



**Figure 8**

Une politique est le reflet d'un consensus entre les parties.

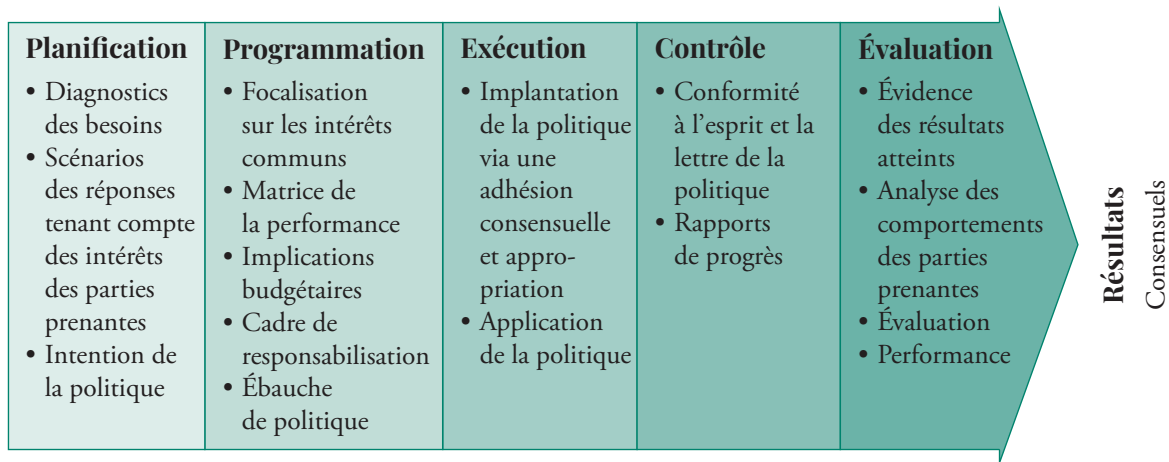
La caractéristique la plus marquée du niveau des résultats-optimisation de développement est certainement la constante négociation entre les parties prenantes, comme en témoigne le schéma du processus de gestion d'une politique ci-après. Par exemple, dans le cas d'un changement de gouvernement, ou d'une prise de conscience de la population sur la mauvaise distribution des bénéfices socio-économiques, l'équilibre du consensus pourrait être compromis.



**Figure 9**

Niveau de gestion et d'analyse évaluative de l'optimisation des résultats de développement.

Il existe toutefois un processus de bonne gestion de l'articulation, de l'application et du suivi d'une politique, présenté dans le schéma ci-dessous.



**Figure 10**

Niveau de gestion et d'analyse évaluative de l'optimisation des résultats de développement.

À ce sixième niveau, les efforts d'évaluation portent sur les considérations de l'atteinte des objectifs, la pertinence, la cohérence, l'efficacité et l'efficience (République française, 2020).



## Conclusion

Au début du chapitre, nous avons vu qu'il existe une pléthore d'objets d'évaluation. Un recensement des interventions de développement international nous a conduits à identifier six niveaux d'objets d'évaluation des interventions de développement. Il nous est possible de conclure que tous les éléments de la gestion du développement international peuvent être évalués et que les approches et les outils pour le faire existent. Chacun des six niveaux d'objets d'évaluation peut faire l'objet d'un examen singulier ou permettre une réflexion qui couvre plusieurs niveaux. L'expérience nous a démontré que chaque niveau d'objets d'évaluation a son approche et ses outils répondant à ses finalités évaluatives. En effet, il faut éviter de croire que l'évaluation d'une politique pourrait être réalisée en utilisant l'outillage pertinent à l'évaluation de la gestion d'un projet ou en analysant une grappe de projets. L'évaluation de politique a sa propre démarche évaluative qui juge de la qualité de la politique concernée en utilisant des outils spécifiques à la gestion du cycle de politique de développement. Alors que tous les éléments du développement international peuvent faire l'objet d'évaluation, il faut être vigilant et utiliser les approches méthodologiques propres à chacun des niveaux d'objets d'évaluation.

## Exercice de réflexion

Déterminer le niveau des objets d'évaluation des situations suivantes :

1. Évaluation d'une cible du développement durable.
2. Évaluation d'un programme de formation récemment donné.
3. Évaluation des effets de la campagne d'immunisation réclamée par la ministre.
4. Évaluation d'une intervention internationale sur la libération de 800 enfants-soldats pour tirer des enseignements de sa réussite.
5. Évaluation d'une catastrophe écologique.
6. Évaluation d'un projet achevé de pont reliant deux régions pour en connaître les retombées socio-économiques.
7. Évaluation d'un décret gouvernemental sur le port d'une ceinture de sécurité dans tout véhicule. Le gouvernement a émis un décret obligeant le port d'un ceinturon de sécurité dans tout véhicule en mouvement.

## Références bibliographiques

- Barbier, J-M. (1985). *L'évaluation en formation*. Presses universitaires de France.
- Beaudry, J. et Gauthier, B. (1992). L'évaluation de programme. Dans B. Gauthier (dir.), *Recherche sociale : de la problématique à la collecte de données* (2<sup>e</sup> éd., p. 425-452). Québec : Presses de l'Université du Québec.
- Brouselle, A., Champagne, F., Contandriopoulos, A-P. et Hartz, Z. (2011). *L'évaluation : concepts et méthodes* (2<sup>e</sup> édition). Presses de l'Université de Montréal.
- Frank, F. et Smith, A. (2000). *Guide du partenariat*. Développement des ressources humaines Canada. <https://publications.gc.ca/collections/Collection/MP43-373-1-2000F.pdf>
- Lusthaus, C., Adrien, M-H., Anderson, G., Carden, F. et Montalván, P. (2003). *Évaluation organisationnelle : Cadre pour l'amélioration de la performance*. Les Presses de l'Université Laval.
- Meier, O. (2017). Frederick Winslow Taylor – Le management scientifiques des entreprises. Dans I. Huault et S.C. Petit (dir.), *Les grands auteurs en management* (3<sup>e</sup> édition, p. 67-79). Éditions EMS.
- Naudet, J-D. (1999). *Vingt ans d'aide au Sabel – Trouver des problèmes aux solutions*. OCDE.
- OCDE. (2005). *Déclaration de Paris sur l'efficacité de l'aide au développement*. OCDE Publishing.
- Osborne, D. et Gaebler, T. (1992). *Reinventing Government. How the Entrepreneurial Spirit is Transforming the Public Sector*. Plume.
- Parsons, T. (1977). *Social Systems and the Evolution of Action Theory*. The Free Press.
- Practical Concepts Inc. (1974). *Guidelines for Teaching Logical Framework Concepts*. [https://pdf.usaid.gov/pdf\\_docs/pnaec576.pdf](https://pdf.usaid.gov/pdf_docs/pnaec576.pdf)
- Project Management Institute. (2021). *Guide du corpus des connaissances en management de projet* (7<sup>e</sup> édition).
- République française. (2020). L'évaluation des politiques publiques : un instrument au service de la réforme de l'État. *Vie publique – République française*. <https://www.vie-publique.fr/eclairage/21987-evaluation-des-politiques-publiques-instrument-de-la-reforme-de-letat>
- Résolution 55/2 de l'Assemblée générale. (2000, 13 septembre). *Déclaration du millénaire*, A/RES/55/2. <http://undocs.org/fr/A/RES/55/2>
- Résolution 70/1 de l'Assemblée générale. (2015, septembre). Transformer notre monde : *le Programme de développement durable à l'horizon 2030*, A/Res/70/1. [https://www.unfpa.org/sites/default/files/resource-pdf/Resolution\\_A\\_RES\\_70\\_1\\_FR.pdf](https://www.unfpa.org/sites/default/files/resource-pdf/Resolution_A_RES_70_1_FR.pdf)
- Sansom, K. (2015). *Introduction au cadre logique*. WEDC – Université Loughborough.
- Taylor, F.W. (2014). *The Principles of Scientific Management*. Martino Fine Books.

# L'identification des parties prenantes en évaluation



*Kris Essosola Eale*  
*Mary Pat Selvaggio*

## Introduction

Une « partie prenante » comporte différentes significations, et il existe plusieurs définitions, parfois contradictoires, de ce terme. Ce chapitre souhaite rendre visibles la place des parties prenantes en évaluation, leurs dynamiques politiques et l'importance de leur identification dans la démarche évaluative. Enfin, il suggère des pistes pour savoir comment les identifier et les hiérarchiser dans leur engagement dans la démarche évaluative.

## Pourquoi les parties prenantes en évaluation ?

Le tableau 1 ci-dessous expose les différentes définitions de « parties prenantes » que l'on peut retrouver dans la littérature. Malgré les divergences dans la définition de ce terme, d'après notre pratique de l'évaluation, la plupart des gens conviendraient qu'étant donné les intérêts des parties prenantes, elles devraient être associées à la prise de décision concernant toutes les interventions qui les concerne, y compris l'évaluation de ces interventions. Ceci assurerait que les résultats de l'évaluation soient significatifs et utiles pour améliorer ces interventions.

Patton et Campbell-Patton (2021), dans leur travail sur l'évaluation axée sur l'utilisation (UFE), associent les principales parties prenantes à des « utilisateurs de l'évaluation » potentiels, et soulignent que « les utilisateurs prévus de l'évaluation sont plus susceptibles d'utiliser les résultats de l'évaluation s'ils comprennent et se sentent responsables du processus et ils sont susceptibles de comprendre et de se sentir propriétaires des résultats de l'évaluation s'ils ont été activement impliqués » ([traduction libre] Patton et Campbell-Patton, 2021). Cullen note que l'inclusion des parties prenantes augmente à la fois le respect de certains principes éthiques d'évaluation, dont notamment l'équité et l'inclusivité, et renforce la crédibilité et la fiabilité de la démarche évaluative, incluant les résultats de l'évaluation (Cullen, 2009, p. 3).

**Tableau 1 — Qui est une partie prenante**

Définitions générales	
Weiss, Carol H.	« Les parties prenantes désignent soit les membres des groupes qui sont manifestement touchés par l'intervention et qui seront donc vraisemblablement affectés par les conclusions de l'évaluation sur l'intervention, soit les membres des groupes qui prennent des décisions sur l'avenir du programme, comme les décisions de continuer ou d'interrompre le financement ou modifier les modes de fonctionnement du programme » (Weiss, 1993, traduit librement par les auteurs).
Banque mondiale	« Les personnes affectées par le résultat – négativement ou positivement, ou ceux qui peuvent influencer sur le résultat d'une intervention proposée » (World Bank, 1996, traduit librement par les auteurs).
Project Management Institute (Larry W. Smith)	« Les individus et les organisations qui sont activement impliqués dans le projet, ou dont les intérêts peuvent être affectés positivement ou négativement à la suite de l'exécution ou de la réussite du projet » (Smith, 2000, traduit librement par les auteurs).
Cullen, Anne (University of Michigan)	« Les parties prenantes désignent toutes les personnes liées à un programme, y compris celles qui ont financé, conçu, mis en œuvre et qui ont été touchées (à la fois positivement et négativement) par l'intervention » (Cullen, 2009, traduit librement par les auteurs).
Project Management Institute (Le guide de la norme ISO 21500)	« Une personne, un groupe de personnes ou une organisation qui a un impact ou pourrait être touché par l'intervention » (Project Management Institute, 2017).
Landau, Peter	« Une partie prenante est un individu, un groupe ou une organisation qui sont touchés par le résultat d'un projet. Ils ont un intérêt dans la réussite du projet et peuvent être à l'intérieur ou à l'extérieur de l'organisation qui parraine le projet » (Landau, 2019, traduit librement par les auteurs).
Définitions propres à l'évaluation	
Centres de contrôle et de prévention des maladies, Bureau de l'exécution et de l'évaluation des programmes (PPEO)	« Personnes ou organisations investies dans l'intervention, intéressées par les résultats de l'évaluation et/ou ayant un intérêt dans ce qui sera fait avec les résultats de l'évaluation » (PPEO, 2012, traduit librement par les auteurs).
Patton et Campbell-Patton	« Les principaux utilisateurs prévus des résultats de l'évaluation; ceux qui ont un intérêt direct et identifiable dans l'évaluation et son utilisation » (Patton et Campbell-Patton, 2021, p. 22, traduit librement par les auteurs).
Better Evaluation	« Les personnes ou les organisations qui seront affectées de manière significative par les résultats du processus d'évaluation ou qui sont affectées par la performance de l'intervention, ou les deux » (Better Evaluation, traduit librement par les auteurs).

Définitions propres à l'évaluation (suite)	
Marceau et Sylvain, dans le chapitre 2 : <i>Terminologie de l'évaluation</i>	<p>« Acteur, actrice qui est impliqué dans la conception ou la mise en œuvre d'une intervention ou qui est susceptible d'être touché par les effets d'une intervention. »</p> <p>La définition comprend le terme « acteur, actrice » comme définisseur initial qui correspond aux « personnes physiques ou morales qui jouent un rôle dans une intervention. ». Les « acteurs, actrices » sont définis en fonction de leur position par rapport à l'intervention : « autorité », « partie prenante », « population cible », « maître d'œuvre », « population exposée », « société » entre autres; avec un vocabulaire provenant de l'analyse des politiques publiques plutôt que de la coopération internationale.</p> <p>Voir le chapitre « Terminologie de l'évaluation » dans cet ouvrage pour plus de détails<sup>1</sup>.</p>

## Les dynamiques politiques et l'importance de l'identification des parties prenantes dans la démarche évaluative

Au cours de nos longues années de pratique de l'évaluation, nous avons observé l'utilisation d'une définition très étroite des parties prenantes de l'évaluation. Cette définition est le plus souvent limitée aux représentants des organisations de mise en œuvre et/ou des organismes de financement, et parfois à une petite branche du gouvernement. Cullen note qu'une telle orientation étroite exclut de nombreuses autres parties prenantes légitimes et pertinentes qui pourraient et devraient être engagées dans la démarche évaluative (Cullen, 2009). Une telle focalisation étroite est en partie due à un biais inhérent des organisations qui commandent les évaluations (généralement les organisations qui financent ou mettent en œuvre les interventions), où la portée de l'évaluation est principalement axée sur leurs besoins. Nous constatons cette dynamique lors de l'examen des termes de référence des évaluations, qui exigent rarement que d'autres parties prenantes soient incluses dans la portée de l'évaluation, selon nous, probablement pour des raisons budgétaires, pour des raisons d'accessibilité ou pour d'autres raisons (inconnues). Pourtant, les évaluateurs sont confrontés à des défis dans l'identification et la sélection des parties prenantes à impliquer dans la démarche évaluative, en partie en raison des lacunes programmatiques (telles que la faible implication des parties prenantes dans les interventions, ainsi que la mauvaise documentation), mais aussi en raison de la dynamique de pouvoir entre les parties prenantes – où les acteurs plus puissants éclipsent la participation et la contribution des parties prenantes « plus faibles ». En effet, Cullen note que « tout comme les programmes qu'ils évaluent, les évaluations sont influencées par de nombreuses décisions politiques » (Cullen, 2009, p. 2) affectant la sélection des évaluateurs, les questions d'évaluation, la conception, l'approche, la mise en œuvre et la diffusion des résultats.

À ce titre, les évaluateurs doivent être conscients de la manière dont les forces politiques influencent le choix de l'approche d'évaluation. Quelles parties prenantes (le cas échéant) participent à l'évaluation? À quelle étape du processus d'évaluation les parties prenantes participent-elles? Qui est

1. De notre côté, ce chapitre et celui sur l'implication des parties prenantes reprend les termes usuels utilisés dans le domaine de la coopération internationale et en Afrique. Il appartient au lecteur de faire la part des choses concernant l'usage de termes plutôt que d'autres en fonction de ce qui a du sens dans sa pratique.

responsable de la prise de décision en matière d'évaluation ? À la lumière de la littérature explorée et de notre expérience sur le terrain, comprendre les influences politiques autour de l'engagement des parties prenantes dans les évaluations aidera les évaluateurs à planifier et/ou protéger :

1. La crédibilité et la fiabilité des résultats;
2. L'utilité, l'utilisation des résultats et la mise en œuvre des recommandations, le cas échéant;
3. La diversité dans la participation des parties prenantes et tirer des conclusions plus inclusives;
4. L'allocation du temps et des ressources financières dans la participation des parties prenantes;
5. Le changement social/dynamique;
6. Les difficultés attachées aux limites des compétences techniques des parties prenantes;
7. L'autonomisation/renforcement des capacités des parties prenantes les moins influentes.

Pour renforcer la crédibilité et la fiabilité de toute évaluation, l'engagement des parties prenantes doit également être inclusif et éthique. Il incombe donc aux évaluateurs de plaider pour l'inclusion de toutes les parties prenantes concernées dans la mesure du possible.

## **Identification des parties prenantes de l'évaluation : classification/typologie**

L'identification des parties prenantes est la première étape importante pour le succès de l'évaluation, et elle commence dès la conception de l'intervention. On ne peut donc pas parler de « parties prenantes de l'évaluation » sans souligner à quel point il est important, pour les évaluateurs, de retracer et de documenter toutes les parties prenantes de l'intervention. En outre, Ramírez et Brodhead ainsi que Patton et Campbell-Patton soulignent que les « utilisateurs » de l'évaluation doivent être identifiés dès le début de l'évaluation (Ramírez et Brodhead, 2013; Patton et Campbell-Patton, 2021).

Par conséquent, l'identification des parties prenantes dans les évaluations est en grande partie le produit de leur rôle, de leur niveau d'implication, de leur sensibilisation, de leur intérêt et de leur influence dans la mise en œuvre de l'intervention, ainsi que de leur besoin d'utiliser les résultats de l'évaluation. Le bureau de la performance et de l'évaluation des programmes du *Center for Disease Control and Prevention* des États-Unis (CDC-PPEO) recommande également d'identifier les parties prenantes en fonction de leurs rôles et de leur contribution à l'évaluation en répondant aux questions suivantes (Salabarría-Peña *et al.*, 2012) :

1. Qui est concerné par l'intervention ?
2. Qui est impliqué dans les opérations de l'intervention ?
3. Qui utilisera les résultats de l'évaluation ?

4. La participation des parties prenantes sélectionnées :
  - a. Augmenterait-elle la crédibilité de l'évaluation?
  - b. Contribuerait-elle à éclairer la mise en œuvre des interventions qui sont au cœur de cette évaluation?
  - c. Contribuerait-elle à plaider en faveur de changements pour institutionnaliser les résultats de l'évaluation?
  - d. Permettrait-elle de mobiliser des fonds et d'autoriser la poursuite ou l'extension de l'intervention?

Abondant dans le même sens, Better Evaluation note que différentes parties prenantes peuvent être engagées à des fins différentes et à différentes phases ou étapes de l'évaluation (Better Evaluation, s. d.). Nous notons qu'il est important que chaque groupe de parties prenantes identifié soit représenté par des personnes intéressées, bien informées, ouvertes, connectées à un groupe important de parties prenantes, crédibles, aptes à apprendre et disponibles pour une interaction tout au long du processus d'évaluation.

Finalement, les évaluateurs doivent établir une classification/typologie des parties prenantes en utilisant une cartographie. Celle-ci permettra de déterminer la participation des parties prenantes dans l'intervention ainsi que le degré de leur participation à la démarche évaluative. Le CDC-PPEO note trois grands groupes de parties prenantes dans les évaluations par rapport à leur rôle dans l'intervention et l'utilisation des résultats de l'évaluation (Program Performance and Evaluation Office, 2012) :

1. **Les personnes impliquées dans les opérations/activités de l'intervention** : la direction, le personnel de l'intervention, les partenaires, les agences de financement (partenaires techniques et financiers) et les membres de la coalition.
2. **Personnes desservies ou affectées par l'intervention** : Bénéficiaires, associations ou organismes impliqués dans des activités de plaidoyer, membres de la communauté et élus locaux.
3. **Les utilisateurs prévus des résultats de l'évaluation** : les personnes en mesure de prendre des décisions concernant l'intervention, telles que les partenaires, les organismes de financement, les membres des associations et le grand public ou les contribuables.

La classification/typologie que nous proposons ci-dessous présente trois catégories de parties prenantes et leur intérêt, leur rôle, leurs responsabilités et leur influence sur les résultats de l'intervention et de l'évaluation. Ces trois groupes sont examinés plus en détail ci-dessous.



**Tableau 2 — La classification/typologie des parties prenantes. À partir de l'expérience des auteurs et de Marchand, 2018, Patton, 2008, PPEO, 2012 et World Bank, 1996.**

Terme	Catégorie	Sous-catégorie	Définition	Intérêt/influence de leur participation au processus d'évaluation et de l'intervention	Facteurs influençant leur participation au processus d'évaluation
Bénéficiaires	Bénéficiaires directs	Population cible	Personnes visées par l'intervention pour modifier leur comportement et/ou améliorer leur bien-être.	<ul style="list-style-type: none"> <li>▪ <b>Accroître la crédibilité de l'évaluation</b> : En tant que principaux bénéficiaires des activités de l'intervention, qui d'autre peut fournir un meilleur retour d'information sur l'efficacité de l'intervention?</li> <li>▪ <b>Déterminer la pertinence des activités de l'intervention</b> : Elles peuvent fournir des informations importantes sur l'efficacité de l'intervention et sur ce qui est plus approprié et pertinent.</li> </ul>	<ul style="list-style-type: none"> <li>▪ <b>Niveau de sensibilisation/connaissance</b> de l'intervention : il permet aux parties prenantes de contribuer efficacement et de manière significative au processus d'évaluation.</li> <li>▪ <b>Capacité des bénéficiaires</b></li> <li>▪ <b>Budget d'évaluation</b></li> </ul>
		Bénéficiaires indirects	Autres personnes/organisations affectées positivement ou négativement par l'intervention.	<ul style="list-style-type: none"> <li>▪ <b>Plaider ou autoriser les changements</b> à l'intervention que l'évaluation peut recommander.</li> <li>▪ <b>Aider à diffuser les résultats de l'évaluation et influencer les politiques des pays</b> à adopter les recommandations de l'évaluation.</li> <li>▪ <b>Apprendre de l'évaluation</b> pour améliorer les interventions.</li> </ul>	<ul style="list-style-type: none"> <li>▪ <b>Niveau de sensibilisation/connaissance</b> de l'intervention.</li> <li>▪ <b>Niveau d'implication dans la mise en œuvre</b> de l'intervention.</li> <li>▪ <b>Identification d'une personne-ressource</b> directement impliquée dans l'intervention.</li> <li>▪ Budget d'évaluation.</li> </ul>
	Les financeurs/donateurs/partenaires techniques et financiers	<p>Les fonds monétaires et banques</p> <p>Les agences internationales et gouvernementales</p> <p>Les organisations philanthropiques et les fondations</p>	Ce sont les organisations qui contribuent au financement de l'intervention.	<ul style="list-style-type: none"> <li>▪ <b>Prendre des décisions sur la poursuite ou l'expansion</b> de l'intervention en fonction des résultats de l'évaluation.</li> <li>▪ Préconiser ou autoriser les changements à l'intervention que l'évaluation peut recommander.</li> <li>▪ Aider à diffuser les résultats de l'évaluation et influencer les politiques des pays à adopter les recommandations de l'évaluation.</li> <li>▪ Apprendre de l'évaluation pour améliorer l'intervention.</li> </ul>	<ul style="list-style-type: none"> <li>▪ <b>Disponibilité des personnes</b> directement impliquées dans la mise en œuvre de l'intervention.</li> <li>▪ <b>Identification d'une personne-ressource</b> dédiée au processus de l'évaluation.</li> <li>▪ Le <b>pouvoir des représentants des organisations financières</b>, très souvent, fait obstacle à la liberté d'expression des autres parties prenantes lors des interactions.</li> </ul>
Organisation de mise en œuvre	Organisation récipiendaire, exécutants	Réceptaire principal Les sous-réceptaires	Ce sont les organismes qui reçoivent les fonds et mettent en œuvre l'intervention. Nous pouvons citer les organisations non gouvernementales (ONG) locales et internationales, les universités, etc.	<ul style="list-style-type: none"> <li>▪ <b>Plaider auprès des gouvernements ou autoriser les changements</b> que l'évaluation recommande.</li> <li>▪ <b>Aider à diffuser les résultats de l'évaluation et influencer les politiques</b> pour adopter les recommandations de l'évaluation.</li> <li>▪ <b>Apprendre de l'évaluation</b> pour améliorer l'intervention.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Dépositaire des données et connaissances de l'intervention.</li> <li>▪ Identification d'une personne clé pour faciliter l'accès des évaluateurs aux données et aux personnes-ressources.</li> </ul>



## Les bénéficiaires

Les définitions du terme bénéficiaire, comme les parties prenantes, varient considérablement. La définition de la FAO (2005) est la plus large – et inclut toute personne ou tout organisme qui tire un certain bénéfice de l'intervention (RuralInvest, 2005), par exemple : les populations cibles de l'intervention, leurs familles, les communautés, les organisations de la société civile locale opérant dans le secteur et la région qui bénéficie des retombés de l'intervention, ainsi que les autorités locales et nationales. Alors que le Centre d'excellence de la gestion axée sur les résultats du Canada définit un bénéficiaire comme un « groupe de personnes qui vivent le changement d'état, de condition ou de bien-être au niveau des résultats ultimes d'un modèle logique » (Affaires mondiales Canada, 2016, p. 9). Cette définition exclut les gouvernements qui sont classifiés comme intermédiaires.

La plupart des publications, y compris celles du Programme des Nations Unies pour le développement (PNUD), CDC et Vision Mondiale, reconnaissent deux types de bénéficiaires :

1. Les bénéficiaires directs, également appelés bénéficiaires primaires;
2. Les bénéficiaires indirects, également appelés bénéficiaires secondaires.

« **Intermédiaire** : Personne, groupe, institution ou gouvernement qui n'est pas le bénéficiaire ultime du projet, mais qui vivra un changement de capacité (résultat immédiat) et un changement de comportement, de pratiques ou de rendement (résultat intermédiaire) qui lui permettront de contribuer à la réalisation d'un changement d'état durable (résultat ultime) des bénéficiaires. » (Affaires mondiales Canada, 2016, p. 9)

Marchand définit les bénéficiaires directs ou primaires comme des personnes spécifiquement ciblées par l'intervention pour changer leur comportement et/ou améliorer leur bien-être (Marchand, 2018). Le manuel de suivi de l'impact du Partenariat Vital Voice Global (*Vital Voices Global Partnership*, VVGP) définit les bénéficiaires directs (ou principaux) comme ceux qui en retirent un bénéfice, c'est-à-dire les individus dont le bénéfice ou le service peut être directement attribué à l'intervention (Vital Voices Global Partnership, 2017, p. 35). Ces deux définitions sont plus génériques et excluent certaines personnes qui ne sont pas ciblées par l'intervention, mais qui peuvent être considérées comme bénéficiaires directes, selon l'exemple dans l'encadré ci-dessous que nous proposons. Le Centre d'excellence de la gestion axée sur les résultats du Canada est encore plus restrictif que Marchand et VVGP dans sa définition de bénéficiaire, en ne considérant qu'un seul type de bénéficiaire dit « bénéficiaire ultime » et classant les autres bénéficiaires comme intermédiaire. De son côté, Vision Mondiale propose une définition plus ouverte des bénéficiaires directs en tant que personnes ou organisations directement touchées par l'intervention, qu'elle soit intentionnelle ou non (World Vision : Strategy and Evidence Unit, 2015). Cette définition est plus inclusive et implique ceux qui ne sont pas intentionnellement visés par l'intervention, mais qui en bénéficient directement; par exemple, les tuteurs et les familles des orphelins et des enfants vulnérables (OEV) sont considérés comme des bénéficiaires directs, alors que l'intervention cible principalement les OEV. Dans le cadre de cet ouvrage, nous favorisons l'inclusion comme Vision Mondiale, nous définissons les bénéficiaires directs comme des individus et/ou des institutions ciblées ou non par l'intervention, mais qui bénéficient directement de l'intervention, qu'elle soit intentionnelle ou non. Ceci a l'avantage de stimuler un sentiment d'appropriation et de favoriser la durabilité des interventions.

Les bénéficiaires indirects ou secondaires, par contre, sont définis comme d'autres personnes ou organisations qui ne sont pas directement ciblées par l'intervention, mais qui en bénéficient encore d'une manière ou d'une autre. Il peut s'agir d'autres membres de la communauté, ou des organisations locales de la société civile opérant dans la région et le secteur. L'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO) note que les bénéficiaires indirects sont souvent, mais pas toujours, ceux qui vivent dans la zone d'influence de l'intervention (RuralInvest, 2005). VVGP note que les bénéficiaires indirects ou secondaires sont des personnes dont le bénéfice ou le service peut être attribué à un bénéficiaire direct de l'intervention mise en œuvre par l'organisation qui fait l'objet du suivi (Vital Voices Global Partnership, 2017, p. 37).

### ***Exemple de bénéficiaires indirects de notre pratique d'évaluation***

*L'intervention contre la tuberculose (TB) dans le secteur minier (TIMS) est mise en œuvre par le Wits Health Consortium dans dix (10) pays d'Afrique australe. Les bénéficiaires directs du programme sont les mineurs, les anciens mineurs et leurs familles, tandis que les bénéficiaires indirects sont les associations de mineurs (organisations locales de la société civile des mineurs), les communautés vivant dans les zones où les interventions sont mises en œuvre et les programmes nationaux de lutte contre la tuberculose dans chaque pays. Dans le cadre de cette intervention, la TIMS a mis en place des Centres de Services de Santé au Travail (CSST) qui sont de petites cliniques spécialisées pour la tuberculose dans les régions minières. Alors même si le TIMS ciblait les mineurs, les anciens mineurs et leurs familles en tant que principale population touchée par la tuberculose, la communauté locale, la société civile (associations de mineurs) et les programmes nationaux de lutte contre la tuberculose de chaque pays ont également bénéficié de la prestation de services des CSST et du renforcement du système de santé national, car les interventions ont renforcé la capacité de leur programme national de lutte contre la tuberculose à faire face à cette crise.*

Étant donné que peu d'interventions engagent effectivement les bénéficiaires lors de la conception de leurs interventions, les évaluations peuvent fournir aux bénéficiaires (directs et indirects) une voix puissante pour exprimer leur point de vue sur les interventions. Les perspectives des bénéficiaires peuvent offrir des informations précieuses sur la manière dont l'intervention peut être soutenue au-delà de l'aide des bailleurs de fonds.

En outre, la participation des bénéficiaires au processus d'évaluation peut :

1. **Augmenter la crédibilité de l'évaluation** : Les bénéficiaires directs sont mieux placés que toute autre partie prenante pour dire si l'intervention a vraiment eu lieu (principaux témoins de l'intervention) et pour fournir des commentaires sur l'efficacité, sur la pertinence et pour juger si l'intervention est appropriée pour régler les problèmes rencontrés par les bénéficiaires eux-mêmes, ou encore suggérer des façons d'améliorer l'intervention. Fischer note que les bénéficiaires permettent la génération d'informations précieuses, authentiques et exactes qui sont importantes pour l'interprétation de toutes les autres données existantes (Fischler, 2013, p. 20).

2. **Plaider pour des changements à l'intervention, ou les autoriser, à la suite des recommandations formulées lors de l'évaluation** : Dans ce cas, le gouvernement et les organisations de la société civile en tant que bénéficiaires indirects peuvent jouer un rôle majeur en plaidant pour des changements ou en permettant (autorisant) la mise en œuvre des changements recommandés par l'évaluation en particulier, s'ils y ont contribué et qu'ils s'identifient à ces recommandations.
3. **Aider à diffuser les résultats de l'évaluation et influencer les politiques des pays à adopter les recommandations de l'évaluation** : Là encore, le gouvernement et les organisations de la société civile en tant que bénéficiaires indirects peuvent aider à diffuser davantage les résultats de l'évaluation et à étendre leur utilisation dans d'autres domaines du pays ainsi qu'à les utiliser pour informer et influencer les politiques du pays.

### Organisations de financement

Aussi appelées organismes fiduciaires, subventionnaires, partenaires techniques et financiers ou encore bailleurs de fonds, ces organisations financent des interventions (sous forme de programmes ou de projets, voir le chapitre sur les objets d'évaluation). Elles se déclinent en plusieurs types, des organismes à but non lucratif, des fonds monétaires et des banques, des agences internationales et gouvernementales, des organisations philanthropiques et des fondations.

Shapiro note que la plupart des organismes de financement veulent (Shapiro, 2013, p. 7) :

- **Créer un impact ou une différence**—Ils veulent que leur argent compte, ils veulent que le travail qu'ils financent soit une réussite, ils veulent être vus comme ayant du succès. À cet égard, ils souhaitent que l'évaluation soutienne et justifie le succès de l'intervention qu'ils financent.
- **Acquérir et partager des informations et des connaissances**—Obtenir une meilleure compréhension de la situation, du problème qu'ils essaient de résoudre, ajouter de la valeur avec leurs interventions choisies aux méthodes existantes et à l'approche utilisée pour résoudre le problème ou la situation existante.
- **Pour accroître leur influence**—Ils veulent accroître leur influence en s'attaquant à ce qu'ils considèrent être les problèmes du monde, du pays ou d'une région particulière.

Par notre expérience, nous notons que faire participer les organismes de financement dans la démarche évaluative peut aider à :

1. **La décision concernant la poursuite, l'extension ou les changements** de l'intervention en fonction des résultats de l'évaluation.
2. **La diffusion des résultats** de l'évaluation pour adopter les recommandations d'évaluation et influencer les politiques.

## Organisations de mise en œuvre

Aussi appelées exécutants, les organisations/partenaires de mise en œuvre sont des organisations qui se voient confier la responsabilité de mettre en œuvre les activités des interventions. Elles allouent des ressources (financières, personnelles et matérielles) pour permettre la mise en œuvre des interventions. Ces organisations peuvent être des entités publiques ou non, telles que les organisations non gouvernementales (ONG) locales et internationales, les universités, les centres et bureaux d'études, etc.

Les partenaires de mise en œuvre peuvent jouer un rôle crucial dans la démarche évaluative étant donné qu'ils sont généralement les principaux gardiens des données et des connaissances nécessaires à l'évaluation. En outre, ils peuvent identifier les personnes clés nécessaires pour faciliter l'accès des évaluateurs aux données et aux informations. Plus important encore, les partenaires d'exécution sont les principaux utilisateurs des résultats de l'évaluation, responsables de la mise en œuvre des recommandations de l'évaluation.

## Conclusion

Malgré les nombreux avantages connus d'impliquer les parties prenantes le plus tôt possible dans le processus d'évaluation, l'identification et la participation des parties prenantes restent un défi pour les évaluateurs. Cela est principalement dû aux contraintes inhérentes au cycle de vie de l'intervention, car il est difficile pour les évaluateurs d'identifier toutes les parties prenantes de l'intervention au stade de l'évaluation. Ces contraintes sont, entre autres : le manque d'implication, le manque de documentation et d'accessibilité à certaines catégories d'acteurs et d'actrices, ainsi que la pression et la dynamique de pouvoir entre les plus influents et les évaluateurs et évaluatrices. Néanmoins, pour préserver la crédibilité de l'évaluation et favoriser son utilisation, il est de la responsabilité des évaluateurs de naviguer entre une position d'objectivité, consciente des dynamiques de pouvoir en jeu, et une position d'ouverture tout au long de la démarche d'évaluation.

## Messages clés

1. La participation active des parties prenantes à l'évaluation renforce celle-ci, et ce, sur plusieurs aspects, comme la crédibilité et l'acceptation de ces résultats, ainsi que le caractère inclusif de l'évaluation (représentation de la diversité des parties prenantes, y compris des groupes les plus faibles et/ou marginalisés).
2. Les parties prenantes représentent une variété d'acteurs, et leur inclusion peut dépendre d'influences politiques dont l'évaluateur doit être conscient. L'évaluateur doit donc identifier les utilisateurs de l'évaluation, ceux qui auront un pouvoir d'influence important pour favoriser la prise en compte de l'évaluation et la mise en œuvre des recommandations, le cas échéant.
3. L'identification des parties prenantes devrait faire partie du processus de planification de l'évaluation, notamment au niveau de leur(s) rôle(s), intérêts et capacité d'influence. En ce sens, une cartographie des parties prenantes est recommandée (elle demeure un instrument de gestion de l'évaluation, et non pas un produit de l'évaluation).

## Références bibliographiques

- Affaires mondiales Canada. (2016). *La gestion axée sur les résultats appliquée aux programmes d'aide internationale : un guide pratique* (2<sup>ème</sup> édition). Centre d'excellence de la gestion axée sur les résultats. [https://www.international.gc.ca/world-monde/assets/pdfs/funding-financement/results\\_based\\_management-gestion\\_axee\\_resultats-guide-fr.pdf](https://www.international.gc.ca/world-monde/assets/pdfs/funding-financement/results_based_management-gestion_axee_resultats-guide-fr.pdf)
- Better Evaluation. (2013). *Déterminer de quelle façon les décisions concernant l'évaluation seront prises*. [https://www.betterevaluation.org/fr/commissioners\\_guide\\_fr/step1\\_fr](https://www.betterevaluation.org/fr/commissioners_guide_fr/step1_fr)
- Better Evaluation. (s. d). *Understand and engage stakeholders*. [https://www.betterevaluation.org/en/rainbow\\_framework/manage/understand\\_engage\\_stakeholders](https://www.betterevaluation.org/en/rainbow_framework/manage/understand_engage_stakeholders)
- Cullen, A. (2009). *The Politics and Consequences of Stakeholder Participation in International Development Evaluation*. Dissertations, 656, 1-166. <https://scholarworks.wmich.edu/dissertations/656>
- Fischler, M. (2013). *Beneficiary Assessment –a participative approach for impact assessment*. Helvetas Swiss Intercooperation. <https://www.shareweb.ch/site/Poverty-Wellbeing/Documents/Case%20M%20Fischler%20BA%20Central%20America,%20Bolivia,%20Kenya.pdf>
- Landau, P. (2019). *Stakeholder vs. Shareholder : How They're Different & Why It Matters*. Project Manager. <https://www.projectmanager.com/blog/stakeholder-vs-shareholder>
- Marchand, M.-P. (2018). Où en sommes-nous avec l'implication des parties prenantes? *The Canadian Evaluation Society*, 33(2), 189-201. <https://doi.org/10.3138/cjpe.42202>
- Patton, M. Q. (2008). *Utilization-Focused Evaluation*. SAGE Publications.
- Patton, M. Q. et Campbell-Patton, C.E. (2021). *Utilization-Focused Evaluation*. SAGE Publications.
- Program Performance and Evaluation Office (PPEO). (2012). *Program Evaluation : Step 1 : Engage Stakeholders*. Center for Disease Control and Prevention. <https://www.cdc.gov/eval/guide/step1/index.htm>
- Project Management Institute. (2017). *A Guide to the Project Management Body of Knowledge*.
- Ramírez, R. et Brodhead, D. (2013). *Utilization Focused Evaluation : a primer for evaluators*. Southbound. [https://www.researchgate.net/publication/263580952\\_Utilization\\_focused\\_evaluation\\_A\\_primer\\_for\\_evaluators](https://www.researchgate.net/publication/263580952_Utilization_focused_evaluation_A_primer_for_evaluators)
- RuralInvest. (2005). *A participatory approach to identifying and preparing small scale rural investments : preparing and using project profiles*. United Nations Food and Agriculture Organization. <https://www.fao.org/3/a1420e/a1420e00.pdf>
- Shapiro, J. (2013). *Writing a Funding Proposal*. CIVICUS. <https://www.civicus.org/documents/toolkits/Writing%20a%20funding%20proposal.pdf>
- Smith, L.W. (2000). *Stakeholder analysis : a pivotal practice of successful projects*. Project Management Institute Annual Seminars & Symposium [Symposium]. Stakeholder engagement, Texas. <https://www.pmi.org/learning/library/stakeholder-analysis-pivotal-practice-projects-8905>

- Vital Voices Global Partnership. (2017). *Tracking Impact : A Handbook For High Quality Data Collection*. <https://www.awid.org/sites/default/files/atoms/files/handbook-final-single20page-2-11.pdf>
- Weiss, C.H. (1993). Where politics and evaluation research meet. *American Journal of Evaluation*, 14, 93-106. <https://journals.sagepub.com/doi/10.1177/109821409301400119>
- World Bank. (1996). *The World Bank participation sourcebook*. <https://doi.org/10.1596/0-8213-3558-8>
- World Vision : Strategy and Evidence Unit. (2015). *Typology of beneficiaries*. [https://www.wvevidence4change.org/wp-content/uploads/Typology-of-beneficiaries\\_full\\_v1.pdf](https://www.wvevidence4change.org/wp-content/uploads/Typology-of-beneficiaries_full_v1.pdf)

# La gestion d'une évaluation

*Jean Serge Quesnel  
Ada Ocampo*



## Introduction

La gestion d'une évaluation est une démarche évaluative qui est rigoureuse et axée sur les preuves. Les outils de gestion de projet sont utiles à la conduite d'une évaluation. D'abord, il faut tenir compte des facteurs contextuels qui influencent l'intervention. Au cours de la démarche évaluative, l'expérience de la pratique nous a montré qu'il est nécessaire d'être transparent concernant les rôles que jouent les parties prenantes, d'avoir une vision claire des attentes, de favoriser une démarche systématique, d'optimiser des ressources limitées, de produire un rapport basé sur des données fiables qui répondent aux besoins de la prise de décision et d'articuler des constats et des recommandations qui améliorent l'intervention. Les résultats de l'évaluation doivent être disséminés aux fins d'action de suivi et d'apprentissage. Ce chapitre est issu de notre pratique de l'évaluation en contexte de développement, en tant que gestionnaires d'évaluations. Il met en évidence les différentes étapes de la gestion d'une évaluation ainsi que ses éléments clés.

## Gestion d'une évaluation : questionnement préalable et étapes

La gestion d'une évaluation nécessite de savoir qui va mener l'évaluation et qui va prendre des décisions à son propos. BetterEvaluation (s.d.) a établi une série de neuf questionnements qui peuvent orienter la gestion d'une évaluation :

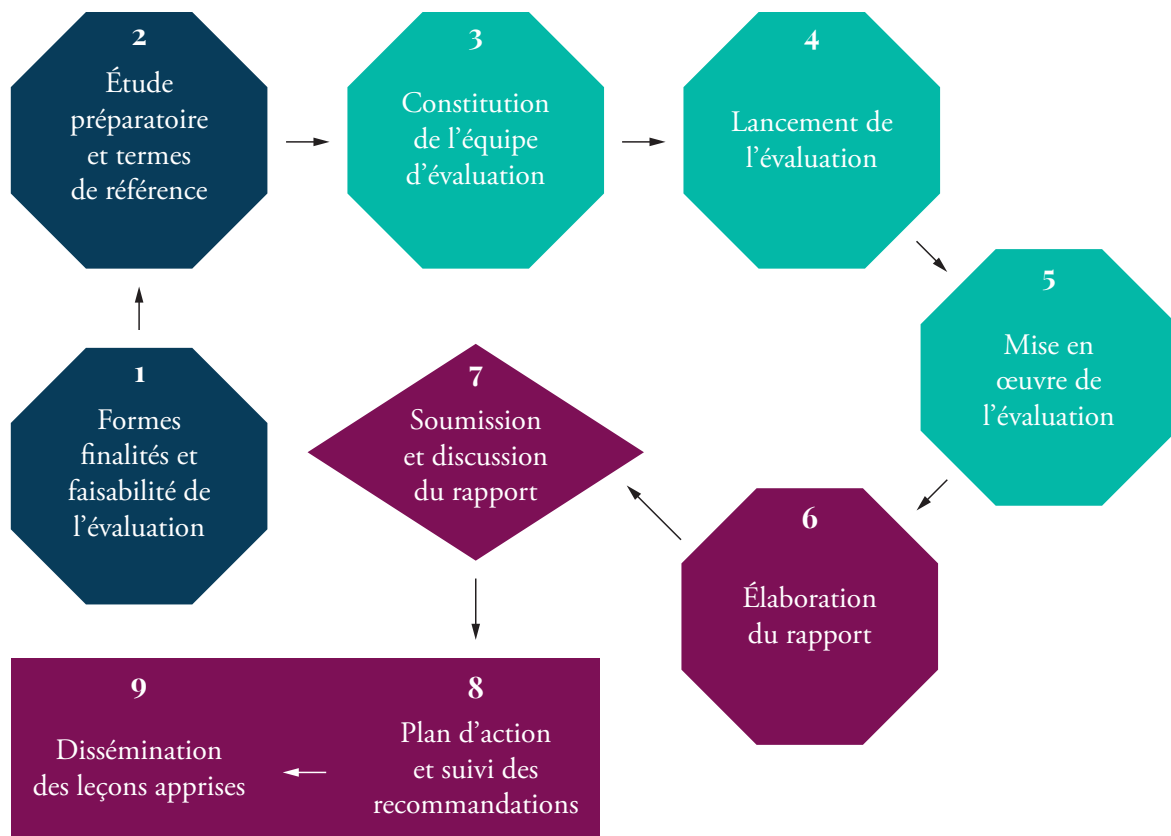
1. Qui va réaliser l'évaluation? S'agit-il d'une autoévaluation, d'une évaluation interne ou externe? Sera-t-elle menée par un employé, un consultant ou par une entité indépendante? S'agit-il d'une évaluation conjointe? Ou participative?
2. Qui a l'autorité pour prendre des décisions concernant l'évaluation, et quels types de décisions? Qui est responsable de conseiller la conduite de l'évaluation? Quel est le processus de prise de décision mis en place pour l'évaluation?
3. Quels acteurs doivent participer à l'évaluation? Comment sont-ils identifiés et impliqués? Deux chapitres dans l'ouvrage traitent plus spécifiquement de ces questions.
4. Quelles sont les ressources financières (disponibles ou à prospecter) et humaines (à la fois internes et externes), et quels sont les délais pour mener l'évaluation?
5. Quels sont les critères qui permettront d'établir que l'évaluation est de qualité et éthique? Comment seront résolus les enjeux éthiques?



6. Comment sera documentée la démarche évaluative? Plus particulièrement les éléments de gestion de projet (par exemple, les types de livrables, le calendrier des livrables, les réunions avec le commanditaire et le principal utilisateur de l'évaluation, les ententes et contrats, etc.)?
7. Quel est le plan de l'évaluation? S'inscrit-il dans un cadre plus général de conduite d'évaluations? Existe-t-il une politique d'évaluation interne à l'organisation commanditaire ou au sein de l'entité indépendante de consultants?
8. Y a-t-il des processus de contrôle de la qualité de l'évaluation, des processus de révision et de validation avant le rapport final?
9. La démarche évaluative envisagée prévoit-elle le renforcement des capacités en évaluation des parties prenantes impliquées? Cette question sera traitée dans le dernier chapitre de l'ouvrage.

La démarche évaluative s'inspire à la fois d'un processus itératif de recherche et d'une démarche de gestion de projet. À l'image de la division de cet ouvrage, un processus de recherche inclura les phases suivantes : une phase de conception, une phase méthodologique, une phase empirique, une phase analytique et une phase de diffusion et de transfert. La démarche évaluative s'inspire généralement de ce processus puisqu'elle emprunte au domaine des sciences économiques et sociales leurs cadres conceptuels, leurs procédés, leurs méthodes, leurs outils de collecte et d'analyse de données. Toutefois, elle se distingue également par la gestion de délais serrés (qui ne sont pas ceux de la recherche), de ressources limitées et de contraintes stratégiques spécifiques. En effet, les questionnements clés de la gestion d'une évaluation montrent notamment que la démarche s'inscrit dans un contexte stratégique de gestion de parties prenantes et de processus de prise de décision qui peuvent être parfois complexes. Ainsi, la figure ci-dessous incorpore des étapes qui sont orientées sur les éléments essentiels de gestion d'une évaluation, et qui reflètent trois phases distinctes simples : la conception, la mise en œuvre et la restitution.





**Figure 1**

Étapes de gestion d'une évaluation.

La suite du chapitre examinera ces étapes sous l'angle spécifique de gestion. De nombreux chapitres dans l'ouvrage examinent des éléments plus spécifiques qui sont inspirés des démarches de recherche.

## Formes, finalités et faisabilité de l'évaluation

Avant de lancer une évaluation, le ou les commanditaires de l'évaluation doivent répondre aux questions suivantes :

**QUOI?** L'objet d'évaluation qui est explicité sous la forme de la modélisation de l'intervention ainsi que sous la forme de questions d'évaluation;

**POURQUOI?** C'est-à-dire la finalité de l'évaluation et les utilisations qui seront faites de l'évaluation par les commanditaires;

**QUI?** C'est-à-dire la forme de l'évaluation et le processus de prises de décision; **COMMENT?** Généralement à travers une matrice d'évaluation (inspirée des approches axées sur les méthodes et les processus de recherche) et le plan d'évaluation.

Une décision qui doit être prise tôt durant la démarche évaluative consiste à déterminer qui réalisera l'évaluation en fonction de quelles finalités. Son corollaire étant celui d'identifier le processus de prise de décision concernant l'évaluation. Les acteurs responsables de la conception et/ou de la mise en œuvre d'une intervention sont généralement ceux qui décident de faire une évaluation (parfois sous la contrainte d'événements, parfois en fonction de processus de reddition de compte) et sous quelle forme ils le feront. Il existe plusieurs formes d'évaluations.

### **Les formes d'évaluations : externe, interne, autoévaluation, conjointe, participative**

Les évaluations externes sont conduites par des services et/ou des personnes extérieures aux acteurs responsables de l'intervention (bailleur de fonds, partenaire de mise en œuvre/exécutant, partenaires gouvernementaux) tandis que les évaluations internes sont menées par un service et/ou des personnes qui dépendent de ces acteurs (OCDE, 2010). Au sein de l'appareil d'État, une évaluation interne impliquerait une équipe d'évaluation interne à un ministère (souvent une direction d'audit, de contrôle et d'évaluation, ou une direction de suivi-évaluation) pour répondre à une commande du ministre, de son cabinet, du secrétaire général, du ministère responsable de l'octroi des crédits ou l'équivalent d'un Secrétariat au Conseil du trésor. Le terme « autoévaluation » fait référence aux évaluations réalisées par ceux qui ont la responsabilité de concevoir et de mettre en œuvre l'intervention (OCDE, 2002). La différence avec une évaluation interne est que ces personnes ne sont pas nécessairement des personnes évaluatrices de métier. Dans une dimension d'autogestion, faire une autoévaluation est la responsabilité de tous. L'évaluation participative implique une conduite d'évaluation grâce à une interaction dynamique entre les parties prenantes où celles-ci ont un pouvoir de décision élevé pour toutes les étapes de la démarche évaluative. Les évaluations conjointes sont des évaluations plus spécifiques au contexte de la coopération internationale et impliquent la collaboration des différents bailleurs de fonds d'une même intervention et leurs partenaires, à des degrés divers d'implication et avec une mise en commun des ressources (OCDE, 2006). L'optique n'est pas celle de l'évaluation participative, mais plutôt celle d'une évaluation entreprise collectivement par plusieurs organismes de coopération pour le développement.

Quelle que soit la forme de l'évaluation, il sera essentiel d'identifier les finalités de celle-ci (sommatrice, formative) ainsi que ce que les parties prenantes souhaitent en faire (l'utilisation de l'évaluation).

### **La finalité et les utilisations possibles d'une évaluation**

Les finalités d'une évaluation sont généralement dénommées par les qualificatifs suivants : sommatives pour les évaluations qui visent à tirer des leçons apprises, et formatives pour les évaluations misant sur l'amélioration de la performance de l'intervention évaluée. Toutefois, ces dénominations ne sont pas suffisantes pour identifier les utilisations potentielles par les parties prenantes, qui peuvent être différentes pour une même évaluation. L'équipe d'évaluation doit être consciente de ces utilisations potentielles, car elles pourront orienter la conception de l'évaluation.

## La faisabilité d'une évaluation

Selon le Secrétariat du Conseil du trésor du Québec (2013), « *une analyse de l'évaluabilité permet de déterminer si une évaluation est appropriée, réalisable et susceptible de fournir une information utile. Elle démontre qu'une intervention peut, ou non, être évaluée, mais également que la conduite d'une évaluation pourra vraisemblablement, ou non, contribuer à répondre aux besoins d'information de son client et améliorer la gestion et les résultats de l'intervention concernée.* »

Selon notre expérience, cinq conditions préalables doivent être considérées avant de réaliser une évaluation :

1. La condition sine qua non demeure la volonté des parties prenantes de réaliser et de contribuer à une évaluation. Sinon, il y aura de la résistance et des embûches.
2. La disponibilité et l'accessibilité des données et des informations qui sont nécessaires pour répondre aux questions d'évaluation. Selon notre expérience, une situation souvent vécue et déroutante, c'est l'existence de données qui ne sont pas fiables.
3. Le ciblage de l'intervention doit être clair. Il est important que l'intention de l'intervention soit sans ambiguïté, mesurable et bien comprise par les parties prenantes. Autrement, il y aura un risque important d'abnégation par certaines parties prenantes avant la mise en œuvre de l'évaluation.
4. L'existence d'un modèle logique de l'intervention aide à expliciter les objectifs et la théorie de changement sous-jacente de l'intervention. Si un tel modèle logique n'existe pas, l'équipe d'évaluation saura en créer un et le faire valider par les parties prenantes.
5. Enfin, il faut un examen réaliste concernant l'enveloppe budgétaire qui est envisagée.

## Étude préparatoire et termes de référence : éléments clés de la conception de l'intervention

Il existe plusieurs façons de concevoir une évaluation. D'après nos expériences, deux types de documents jettent les bases de la portée de l'évaluation.

### L'étude préparatoire

L'étude préparatoire, parfois appelée note méthodologique ou cadre conceptuel, inclut les éléments clés de la conception de l'évaluation. Une bonne étude préparatoire à l'évaluation est indispensable pour déterminer sa finalité et éviter des écueils qui pourraient causer des embarras importants, et même l'échec de l'évaluation. L'étude préparatoire comporte les éléments suivants : la compréhension de l'objet d'évaluation, y compris les motifs pour l'évaluer, la problématique à l'origine de l'intervention, les objectifs/cibles de l'intervention et ses composantes sous forme de modèle logique et/ou théorie de changement, les questions d'évaluation, le choix des approches et leurs méthodologies (c'est-à-dire le choix des méthodes et des outils de collecte et analyse des données, ainsi que leurs limites). Un outil clé de l'étude préparatoire est la matrice d'évaluation qui résume l'ensemble de ces

éléments. L'élaboration d'une étude préparatoire permet de définir l'ampleur de l'évaluation en identifiant l'information nécessaire qui devra être collectée, ce qui permet de déterminer les ressources requises et l'échéancier de l'évaluation.

Selon notre expérience, l'élaboration d'une étude préparatoire n'est pas encore un réflexe chez les commanditaires de l'évaluation. Parfois, seuls des termes de référence sont lancés, et l'étude préparatoire est réalisée par la suite avec l'appui de l'équipe d'évaluation choisie. Dans certains cas, l'étude préparatoire a été réalisée avant les termes de référence, mais sert de base à la discussion avec l'équipe d'évaluation qui pourra l'enrichir de ses propres expériences et expertises.

### Les termes de référence

Les termes de références (TDR) donnent une description du travail spécifique à réaliser dans le cadre d'un mandat d'évaluation. Ils constituent le document qui permet au commanditaire d'une évaluation de passer commande à l'équipe d'évaluation, et de lui expliquer ce qu'il souhaite et pourquoi.

Les TDR comportent : les motifs de l'évaluation, l'historique, le contexte, les visées et les résultats attendus de l'intervention évaluée, les principes et l'approche qui guideront l'évaluation, la méthodologie d'évaluation, la composition de l'équipe et les compétences requises, les modalités ou arrangements institutionnels, les livrables, un calendrier des jalons de l'évaluation, le budget et les exigences en matière de rapports. Le tableau qui suit résume les éléments clés des TDR :

**Tableau 1 — Les éléments des termes de référence d'une évaluation. Adapté de l'ENAP, 2021.**

Présentation générale de l'évaluation attendue	Précisions sur le contenu de la commande	Précisions sur les aspects opérationnels de l'évaluation
Contexte de l'évaluation	Questions d'évaluation	Durée / échéancier
Informations sur l'intervention à évaluer	Méthodes ou approches privilégiées	Budget
Attentes du commanditaire/client principal (par rapport aux objectifs et finalités de l'évaluation)	Répondant.e.s et collectes nécessaires	Composition de l'équipe et profil recherché
Principaux intéressés	Attentes sur le rôle de l'équipe d'évaluation	Livrables attendus

Les termes de référence sont fort utiles pour mandater une équipe d'évaluation interne. Ils sont nécessaires pour le lancement de l'appel de soumissions pour le recrutement d'équipes d'évaluation externes.

## La constitution de l'équipe d'évaluation

La première étape de mise en œuvre de l'évaluation consiste à constituer une équipe d'évaluation. Si l'évaluation est conduite par une équipe interne à l'organisation, il revient au responsable de l'unité d'évaluation de constituer l'équipe d'évaluation en nommant un chef d'évaluation, en assignant des professionnels en évaluation au sein de l'équipe et en impliquant des ressources provenant d'autres unités administratives de l'organisation pouvant contribuer à l'évaluation.

Si l'évaluation est conduite par une expertise externe en évaluation, la personne responsable de l'unité d'évaluation de l'organisation commissionnant l'évaluation publiera les termes de référence de l'évaluation et procédera à une demande de propositions pour la réalisation de l'évaluation. Une fois la date limite des soumissions de propositions atteinte, une équipe évaluera avec rigueur les propositions soumises sur la base de critères techniques et financiers. Une négociation sera entamée avec l'équipe externe sélectionnée, ce qui conduira à la signature d'un contrat.

L'expérience nous a démontré que la constitution d'une bonne équipe d'évaluation est primordiale pour la réalisation d'une évaluation de qualité. La compétence des membres d'évaluation et les bonnes relations de travail entre eux favorisent une synergie créatrice au moment de l'analyse des données et de l'élaboration de conclusions basées sur une évidence factuelle. Il faut prévenir les risques associés à la constitution d'une équipe d'évaluation qui pourraient causer des préoccupations majeures. Signalons les suivants :

- a. L'équipe d'évaluation s'avère incompétente : il faudra alors ajouter une expertise complémentaire ou bien arrêter l'évaluation.
- b. L'équipe d'évaluation irrite des parties prenantes : il faut veiller à ce que les premières communications soient bien organisées et que l'équipe d'évaluation soit sensible aux contextes et aux relations interculturelles.
- c. Des membres de l'équipe de l'évaluation ne sont pas discrets : il faut rapidement leur donner un avertissement et rappeler les normes de confidentialité.
- d. L'équipe ne respecte pas les termes de référence : il faut être clair sur le fait que les parties contractuelles doivent se mettre d'accord avant de modifier les TDR ou d'ajouter des éléments à ceux-ci.

## Le lancement de l'évaluation

Le lancement de l'évaluation est une étape préalable à la mise en œuvre qui permet de présenter les membres de l'équipe d'évaluation aux parties prenantes, de s'assurer que l'équipe d'évaluation comprend bien son mandat et qu'elle organise son travail de manière optimale. À cette fin, une réunion de démarrage sera utile pour initier l'équipe d'évaluation à la finalité et au contexte de l'évaluation.

## **Constitution d'un comité d'évaluation**

Il est d'usage de constituer un groupe consultatif d'évaluation, souvent connu comme « comité d'évaluation », qui rassemble des représentants des parties prenantes. Ce groupe a le mandat de donner des avis à l'équipe d'évaluation et de commenter les produits intérimaires générés par l'équipe. Le groupe examinera les ébauches du rapport d'évaluation et donnera des avis sur les constats, les conclusions et les recommandations. Le groupe jouera un rôle important pour faciliter l'accès à l'information et aux contributions des parties prenantes à la démarche évaluative. Il augmente l'acceptabilité et l'utilisation des résultats de l'évaluation et permet de les diffuser plus rapidement et efficacement.

## **Rôles et responsabilités des différents intervenants**

Il y a lieu de clarifier les rôles et responsabilités des intervenants, lesquels seront inclus dans le plan de l'évaluation. Les principaux intervenants sont : le mandataire de l'évaluation, le gestionnaire de l'évaluation, les membres du comité d'évaluation, l'équipe d'évaluation et les experts spécialisés recrutés pour appuyer le travail de l'équipe d'évaluation (enquêteurs, experts techniques), les personnes-ressources d'information, les personnes bénéficiaires de l'intervention et les personnes qui prendront les décisions de suivi à l'évaluation.

## **Le plan de travail de l'évaluation**

À la lumière des informations disponibles et de la réalité du terrain où s'effectue l'intervention évaluée, l'équipe prépare le rapport de démarrage de l'évaluation qui permet une mise à jour des termes de référence. Doté de cette information, le plan de travail de l'évaluation pourra être rédigé. Ce plan contiendra les renseignements suivants :

- Le mandat et l'envergure de l'évaluation;
- Les rôles et responsabilités des parties prenantes;
- Un organigramme;
- Les rôles et responsabilités de chaque personne impliquée;
- L'approche d'évaluation finale;
- Les approches méthodologiques de l'évaluation;
- La description des tâches qui doivent être accomplies;
- La description des exigences professionnelles requises;
- Le coût estimé pour réaliser chaque tâche;
- Le temps requis pour accomplir chaque tâche;
- La description des risques et les façons de les mitiger;
- Un plan de gestion des communications;
- Un tableau qui synthétise tous les livrables de la démarche évaluative;
- Un tableau récapitulatif des allocations budgétaires.

## La mise en œuvre de l'évaluation

Il s'agit maintenant de mettre en œuvre ce qui a été planifié et de réaliser l'évaluation. Conformément au plan de travail de l'évaluation, l'équipe d'évaluation devra organiser les aspects logistiques et le travail de terrain. Les données doivent être collectées à partir des méthodes et outils choisis et auprès des sources identifiées dans la matrice d'évaluation, présentée dans le plan de travail. À tout moment, en fonction des défis rencontrés, on saura adapter la stratégie d'évaluation. Au cours de la mise en œuvre, l'équipe d'évaluation doit maintenir le contact avec les parties prenantes à travers le comité d'évaluation pour fournir ou recevoir des rétroactions et assurer un suivi. Elle doit également s'assurer de l'accord et de l'adhésion du comité d'évaluation lorsque des changements au plan de travail doivent être apportés.

L'effort majeur de la mise en œuvre de l'évaluation est la collecte, le traitement, l'analyse et l'interprétation des données. Généralement, l'équipe d'évaluation procédera à l'examen des données existantes puis à la cueillette de données primaires. Le traitement, l'analyse et l'interprétation des données permettront de répondre aux questions d'évaluation. D'autres chapitres de cet ouvrage sont consacrés à ces activités.

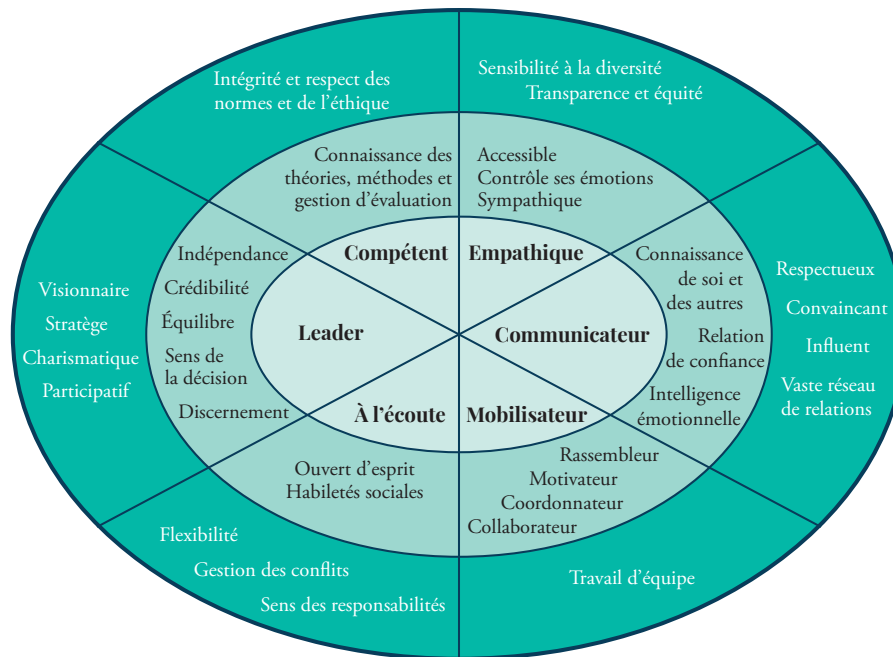
### Gestion de la confidentialité

Il est primordial de garantir aux participants que leurs réponses resteront confidentielles. L'équipe d'évaluation doit donc s'assurer que les résumés d'entrevue ou les réponses aux questionnaires seront gardés dans un lieu sûr selon un processus d'anonymisation des personnes. Il est important que l'identité des participants ne soit pas révélée aux membres du comité d'évaluation ni à qui que ce soit d'autre, pour se protéger contre d'éventuelles conséquences néfastes. La nécessité d'un consentement écrit doit être examinée selon les usages, les règlements des commanditaires, mais aussi selon les règles des organisations auxquelles appartient l'équipe d'évaluation (par exemple, s'il s'agit d'une université, celle-ci est dotée d'un comité d'éthique qui aura un droit de regard sur le respect des politiques institutionnelles) ainsi que des organisations et des personnes qui seront appelées à partager leur opinion.

### Gestion des personnes impliquées dans l'évaluation

La gestion des relations interpersonnelles lors de la démarche d'évaluation est le facteur principal de succès. Créer une atmosphère d'ouverture et de collaboration désamorce les tensions et les craintes, invite à un échange franc et facilite un travail à l'unisson. C'est un rôle clé du gestionnaire de l'évaluation. Ceci prend une importance accrue lorsque l'on collabore en milieux interculturels.

Quand les choses vont bien, c'est un plaisir de travailler en équipe. Lorsque les choses vont mal, les qualités de leadership du gestionnaire sont mises à l'épreuve. Le gestionnaire doit anticiper et mitiger les effets néfastes de nombreux facteurs qui influent sur l'atmosphère existant au sein de l'équipe d'évaluation, tels : les pressions d'un calendrier comprimé, la non-disponibilité de ressources, des goulots d'étranglement dans la livraison de données, la sous-qualité des livrables, la réalisation de suppositions critiques néfastes et l'apparition d'imprévus et enfin, un moral bas. La règle d'or de la gestion des relations interpersonnelles est la franchise et la transparence, le tout dosé d'empathie. Le schéma qui suit présente les compétences clés d'un gestionnaire d'évaluation.



**Figure 2**

Compétences clés d'un gestionnaire d'évaluation.

En bref, la gestion de la mise en œuvre de l'évaluation comprend les tâches qui suivent :

- La réalisation et le suivi du plan de travail de l'évaluation;
- Les rapports périodiques d'avancement des travaux;
- Administration du contrat et des sous-contrats pour assurer des livrables à temps et de qualité;
- Surveillance et amélioration du rendement des personnes impliquées;
- L'exercice de contrôle de la qualité et la rétroaction sur les livrables;
- Les décisions d'accepter ou non les livrables;
- Les décisions de nature logistique;
- Les décisions de couper certaines activités;
- Les décisions de mesures correctives et ajustement au calendrier;
- Le contrôle et les ajustements pour l'optimisation budgétaire;
- Les communications et les relations avec les parties prenantes;
- La gestion des archives.

Une évaluation est un projet et sa réalisation doit être appuyée par les outils de gestion de projet.



## Gestion de l'évaluation malgré les contraintes

Dans son livre *Real world evaluation*, Michael Bamberger examine comment on peut réussir une évaluation tout en tenant compte des nombreuses contraintes qui existent. Ces contraintes sont celles concernant l'information disponible, celles du temps disponible, celles du budget alloué et celles de nature politique. Pour renforcer le devis d'évaluation et sa réalisation, il propose de :

- Bâtir la matrice d'évaluation sur un modèle logique/théorie de changement bien fait et solide; ce faisant, on explicitera la chaîne de causalités et les suppositions. On identifiera aussi les facteurs économiques, politiques, institutionnels, environnementaux et socioculturels afin de pouvoir expliquer les écarts dans la performance et les résultats.
- Adopter une stratégie mixte de collecte et d'analyse des données tout en faisant un usage maximal des données secondaires, y compris les données de suivi de l'intervention ainsi que d'anciennes évaluations.
- Si le calendrier et le budget le permettent, recueillir des données à des points additionnels du cycle de l'intervention.
- Simplifier les instruments de collecte des données.
- Utiliser les données secondaires de façon créative, en exploitant les données de recensements déjà faits pour établir la base des données de la situation initiale et référentielle, et pour les contrôles et les groupes de comparaison.
- Considérer réduire la taille de l'échantillonnage.
- Réduire les coûts de collecte des données en embauchant des personnes qui facturent des honoraires moins élevés pour conduire les entrevues; utiliser l'observation directe au lieu de recensements.
- Travailler avec d'autres évaluations pour la collecte des données.

## La préparation du rapport d'évaluation

La majorité des évaluations comportent comme principal livrable un rapport d'évaluation qui présente la réponse aux questions d'évaluation sous forme de constats. Un autre chapitre de l'ouvrage vous présente les éléments constitutifs d'un rapport d'évaluation ainsi que les différentes façons de présenter les résultats d'évaluation, en dehors d'un rapport écrit. Dépendamment des termes de références, des recommandations ou des pistes d'amélioration peuvent avoir été commandées. Un chapitre de l'ouvrage est consacré à l'élaboration des recommandations pour favoriser leur mise en œuvre.

Neuf critères de mesure de la qualité des rapports d'évaluation sont proposés dans la communauté de pratiques « Capacity4dev » de l'Union européenne :

- *La satisfaction des besoins et des attentes;*
- *La pertinence de l'examen de la raison d'être de l'intervention, de ses réalisations, des résultats et des impacts;*
- *La justification de la méthodologie;*

- *La fiabilité des données;*
- *La solidité de l'analyse;*
- *La crédibilité des constats;*
- *La validité des conclusions;*
- *L'utilité des recommandations;*
- *La clarté du rapport (Evaluation Unit DEVCO, 2015).*

Pour approfondir ses connaissances sur les systèmes institutionnels de contrôle de la qualité des rapports d'évaluation, le lecteur aurait intérêt à se familiariser avec le système institutionnel d'examen de la qualité des rapports d'évaluation du Fonds des Nations Unies pour l'enfance (UNICEF), connu sous l'acronyme GEROS (UNICEF, 2017). La procédure GEROS implique un examen, par une équipe indépendante de consultants, des évaluations réalisées aux niveaux mondial, régional et national, ainsi qu'une méta-analyse annuelle. L'UNICEF a adapté les standards du Groupe des Nations Unies pour l'évaluation (UNEG) comme critères d'analyse de la qualité des rapports.

## **Le contrôle de qualité de la gestion d'une évaluation**

À la fin de la gestion d'une évaluation, le contrôle qualité examine les points suivants :

- L'évaluation a été réalisée conformément à des méthodes et à des processus dont la qualité a été vérifiée. Tout écart par rapport aux méthodes et processus a été corrigé de manière satisfaisante;
- Les données ont été recueillies auprès de sources suffisantes et adéquates, afin de garantir leur crédibilité;
- Les constats reposent sur des analyses valides;
- Les constats, les conclusions et les recommandations sont logiques et cohérents;
- Un nombre suffisant de consultations a été engagé pour veiller à la précision, à la validité, à la pertinence et à l'utilité de l'évaluation;
- L'évaluation tient compte des questions liées aux droits de la personne et à l'égalité des sexes et d'autres normes et principes d'évaluation;
- Les recommandations sont réalistes, prescriptives, susceptibles d'être acceptées et appliquées, et n'entraînent aucune conséquence négative sur des sujets dépassant le champ de l'évaluation;
- Le rapport répond au cahier des charges ainsi qu'à toutes les questions évaluatives;
- Le rapport est rédigé dans un style simple et un format approprié.

### **Plan d'actions de suivi des recommandations**

Un plan d'action de suivi des recommandations doit inclure les éléments suivants :

- Une instance de décision concernant les recommandations à mettre en œuvre, laquelle fera une revue périodique de l'avancement de leur implantation;

- Un support électronique qui facilitera la dissémination du plan de suivi et sa mise à jour;
- Un énoncé clair de l'état de suivi de la mise en œuvre des recommandations, par exemple : recommandation en cours, accomplie, reportée, rejetée;
- Les dates balises;
- Les personnes responsables du suivi.

Ce rapport peut être soumis annuellement aux différentes instances.

## La dissémination des résultats d'évaluation et partage des leçons apprises

Aux fins de mémoire institutionnelle et d'apprentissage, les résultats d'évaluation sont rendus disponibles auprès du personnel et des partenaires des commanditaires. Ceci peut être fait de façon active par des bulletins d'information et des sessions de débriefage. Les résultats peuvent être rendus accessibles par des moyens passifs tel l'accès à une banque électronique des rapports d'évaluation et des réponses de l'administration. Pour un partage plus en profondeur des leçons apprises, une métaévaluation d'un ensemble d'évaluations (internes et externes) peut être réalisée et ainsi permettre de positionner les leçons récentes dans l'ensemble de l'apprentissage institutionnel.

## Conclusion

Pour conclure ce chapitre, nous mettons en évidence les étapes et les principales activités d'une bonne gestion d'une évaluation. Le tableau qui suit les résume.

**Tableau 2 — Étapes, visée et activités de gestion de l'évaluation. Adapté de BetterEvaluation, s. d.**

Étapes		Visées	Activités
1	Formes, finalités et faisabilité	Déterminer comment les décisions concernant l'évaluation seront prises.	<ul style="list-style-type: none"> <li>▪ Identifier les personnes impliquées dans les décisions et expliquer quels seront leurs rôles;</li> <li>▪ Préciser les fonctions du responsable de l'évaluation et des évaluateurs;</li> <li>▪ Traiter des problèmes particuliers de gestion des évaluations liés aux projets conjoints, y compris les partenariats avec les donateurs.</li> </ul>
2a	Étude préparatoire	Définir la portée de l'évaluation.	<ul style="list-style-type: none"> <li>▪ Clarifier ce qui sera évalué;</li> <li>▪ Décrire la théorie du changement;</li> <li>▪ Identifier qui sont les principaux utilisateurs prévus de l'évaluation et établir à quoi elle leur servira;</li> <li>▪ Élaborer des questions d'évaluation clés inclusives;</li> <li>▪ Décider du bon moment pour l'évaluation;</li> <li>▪ Décider si l'évaluation sera faite par une équipe externe, une équipe interne ou un hybride des deux;</li> <li>▪ Déterminer les qualités de l'évaluateur;</li> <li>▪ Identifier les ressources disponibles pour l'évaluation et ce qui sera nécessaire.</li> </ul>

## L'évaluation en contexte de développement

Étapes		Visées	Activités
2b	Termes de référence (TDR)	Élaborer les TDR.	<ul style="list-style-type: none"> <li>▪ Spécifier le contexte de l'évaluation;</li> <li>▪ Préciser le but/objectifs/justification de l'évaluation;</li> <li>▪ Mentionner les utilisateurs et l'utilisation prévue de l'évaluation;</li> <li>▪ Spécifier les questions d'évaluation clés;</li> <li>▪ Indiquer les principes et l'approche qui guideront l'évaluation;</li> <li>▪ Spécifier les méthodologies;</li> <li>▪ Énoncer les rôles et responsabilités des différents acteurs;</li> <li>▪ Préciser les exigences en matière de rapports;</li> <li>▪ Spécifier la chronologie et les jalons de l'évaluation;</li> <li>▪ Souligner toute autre exigence.</li> </ul>
3	Constitution de l'équipe d'évaluation	Constituer l'équipe d'évaluation.	<ul style="list-style-type: none"> <li>▪ Annoncer les termes de référence de l'évaluation (TDR)/la demande de proposition (DP);</li> <li>▪ Sélectionner une équipe d'évaluation;</li> <li>▪ Préparer le contrat;</li> <li>▪ Orienter l'équipe d'évaluation.</li> </ul>
4	Lancement de l'évaluation	Gérer l'élaboration du plan de travail d'évaluation, y compris la logistique.	
5	La mise en œuvre de l'évaluation	Gérer la mise en œuvre de l'évaluation.	<ul style="list-style-type: none"> <li>▪ Formation de l'équipe des agents de terrain;</li> <li>▪ Travail de terrain : collecte de données;</li> <li>▪ Nettoyage des données;</li> <li>▪ Analyse des données (production d'analyses).</li> </ul>
6	Préparation du rapport	Guider la production d'un rapport de qualité.	<ul style="list-style-type: none"> <li>▪ Assurer un contrôle de la qualité du rapport.</li> </ul>
7	Soumission et discussion du rapport	Diffuser les rapports et soutenir l'utilisation de l'évaluation.	<ul style="list-style-type: none"> <li>▪ Préparer des présentations écrites, orales et visuelles adressées aux différents auditoires.</li> </ul>
8	Plan d'action de suivi des recommandations	Systématiser le suivi des recommandations.	<ul style="list-style-type: none"> <li>▪ Préparer un outil de suivi.</li> </ul>
9	Dissémination des leçons apprises	Faire connaître les apprentissages résultant de l'évaluation.	<ul style="list-style-type: none"> <li>▪ Élaborer des outils de dissémination.</li> </ul>

## Exercice de réflexion

On vous demande si vous acceptez de diriger une évaluation d'une durée de six mois, visant à évaluer l'efficacité d'une campagne d'information préventive sur la santé dans votre pays. La campagne consiste en séminaires de deux jours impliquant des professionnels de la santé dans les communautés à travers le pays. La raison de l'évaluation est de déterminer si la campagne d'information a amélioré les pratiques de santé. Vous avez un budget de taille modérée et une équipe de six assistants de recherche qui vous aideront à faire le devis et à réaliser l'évaluation.

1. Selon vous, quelles sont les conditions préalables devant être en place avant de vous engager à diriger cette évaluation? Expliquez votre réponse.
2. Est-ce que cette description du cas est suffisante pour vous permettre de cerner l'envergure de la tâche? Sinon, quels éléments additionnels vous sont nécessaires?
3. Quel est le moyen le plus facile qui vous permettra de clarifier les liens de causalité entre les campagnes d'information et l'amélioration des pratiques de santé?
4. Croyez-vous que cette évaluation soit faisable?
5. Quelles sont les contraintes à surveiller lors de la mise en œuvre de l'évaluation?

## Références bibliographiques

- Bamberger, M. et Mabry, L. (2019). *Real World Evaluation : Working under Budget, Time, Data and Political Constraints* (3<sup>e</sup> édition). SAGE Publications.
- BetterEvaluation (s. d.). *Manage Evaluation*. [https://www.betterevaluation.org/en/rainbow\\_framework/manage\\_evaluation](https://www.betterevaluation.org/en/rainbow_framework/manage_evaluation)
- BetterEvaluation (s. d.). *Étapes du processus d'évaluation*. [https://www.betterevaluation.org/fr/commissioners\\_guide\\_fr](https://www.betterevaluation.org/fr/commissioners_guide_fr)
- École nationale d'administration publique. (2021). *Programme international d'évaluation du développement : Gestion de l'évaluation* [matériel pédagogique, module 7, diapo 8].
- Groupe des Nations Unies pour l'évaluation (UNEG). (2010). *Good Practice Guidelines for Follow up to Evaluations*. <http://www.uneval.org/document/detail/610>
- Groupe des Nations Unies pour l'évaluation (UNEG). (2014). *Intégrer les droits de l'homme et l'égalité des sexes aux évaluations*. <http://www.uneval.org/document/download/1380>
- Groupe des Nations Unies pour l'évaluation (UNEG). (2016). *Normes et règles d'évaluation*. UNEG. <http://www.unevaluation.org/document/download/2700#:~:text=Adopt%C3%A9s%20en%202005%2C%20les%20Normes,Unies%20et%20d'autres%20entit%C3%A9s.&text=Ce%20document%20fon%2D%20dateur%20de,r%C3%A9alis%C3%A9s%20par%20les%20Nations%20Unies>
- OCDE. (2002). *Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats*. <https://www.oecd.org/dac/evaluation/2754804.pdf>
- OCDE. (2006). *Orientations relatives à la gestion des évaluations conjointes*. <https://www.oecd.org/dac/evaluation/dcdndep/37681798.pdf>

- Secrétariat du Conseil du Trésor du Canada. (2006). *Éthique et normes professionnelles pour la collectivité de l'évaluation au gouvernement du Canada*. <https://www.tbs-sct.gc.ca/ceel/career-carriere/pesecgc-enpcegc/pesecgc-enpcegc-fra.asp>
- Secrétariat du Conseil du trésor du Québec. (2013). *Glossaire des termes usuels en mesure de performance et en évaluation. Pour une gestion saine et performante*. [https://www.tresor.gouv.qc.ca/fileadmin/PDF/publications/glossaire\\_termes\\_usuels.pdf](https://www.tresor.gouv.qc.ca/fileadmin/PDF/publications/glossaire_termes_usuels.pdf)
- UNICEF. (2017). *Global Evaluation Report Oversight System : Handbook for UNICEF Staff & Independent Assessors*. <https://www.unicef.org/media/54781/file>
- Union européenne. (s. d.). *Capacity4dev Connecting the Development Community*. <https://europa.eu/capacity4dev/>

# Principes, normes, règles et standards, éthique et déontologie de l'évaluation



*Jean Serge Quesnel*  
*Ahmed Bencheikh*

## Introduction

Dans ce chapitre, nous allons passer en revue la genèse de la normalisation de la fonction d'évaluation et l'état de l'art des normes promues par diverses entités normatives. Nous examinerons les apports d'entités qui ont contribué de façon significative au développement des normes d'évaluation. Enfin, nous allons décrire l'utilité des normes pour faciliter la professionnalisation de la pratique de l'évaluation.

## Genèse des normes

La période du développement moderne de l'évaluation aux fins d'amélioration des décisions publiques remonte à la décennie 1973-1983. La période de la professionnalisation de l'évaluation s'est étalée sur trois décennies (1983–2013) et se poursuit toujours. La création de réseaux et d'associations d'évaluation a été une étape primordiale à la normalisation et à la professionnalisation de l'évaluation. L'établissement de principes, de normes, de règles et de standards a été une étape essentielle pour assurer la fondation de la crédibilité et de la fiabilité de la fonction d'évaluation.

Durant les années 1970-1990 émergent les premières sociétés professionnelles de l'évaluation qui ont organisé des congrès et qui ont fait paraître des publications traitant de l'évaluation. Voici les sociétés qui ont été créées avant l'an 2000 :

- Eastern Evaluation Research Society (1978)
- Société canadienne d'évaluation (1981)
- American Evaluation Association (1986)
- Australasian evaluation society (1987)
- Société québécoise d'évaluation de programme (1988)
- Association africaine d'évaluation (1999)
- Société Française de l'Évaluation (1999)

Depuis l'an 2000, il y a eu une croissance exponentielle du nombre de réseaux et d'associations aux niveaux national, régional et global. L'Organisation pour la coopération en évaluation (OICE) est une entité parapluie qui maintient sur son site Internet un inventaire d'organisations volontaires pour l'évaluation professionnelle (VOPE). En 2021, l'OICE a répertorié 150 VOPE actives.

Une pierre angulaire mise en place en 1981 pour l'établissement de normes et standards d'évaluation a été la contribution du Comité mixte sur les normes d'évaluation en éducation. La deuxième référence constitue les *Principes pour l'évaluation de l'aide au développement*, émis en 1991 par le Comité d'aide au développement (CAD) de l'Organisation pour la coopération et le développement économique (OCDE). Encore aujourd'hui, ces principes sont reconnus mondialement comme l'étalon des normes d'évaluation du développement. La troisième référence mondiale représente les *Normes et règles d'évaluation* produites en 2006 et actualisées dix ans après par le Groupe des Nations Unies pour l'évaluation (UNEG).

Les associations et réseaux régionaux d'évaluation ont aussi produit des normes professionnelles pour l'évaluation. Mentionnons de façon particulière l'Association africaine d'évaluation (AfrEA). L'AfrEA a élaboré en 2002 un guide d'évaluation en Afrique, approuvé par les associations, les réseaux, les évaluatrices et évaluateurs africains. Ce guide constitue un cadre permettant de réaliser des évaluations qui tiennent compte des réalités africaines. Il est conforme aux procédures et aux normes internationales en la matière, et comprend des principes déontologiques et des critères de qualité. Ces principes reflètent un système de valeurs qui sont communes aux acteurs intervenant dans le cadre de l'évaluation en Afrique. L'Association africaine d'évaluation a publié en 2020 la quatrième version des *Principes directeurs pour l'évaluation en Afrique*.

## **Principes, normes, règles et standards, éthique et déontologie pour guider la pratique professionnelle de l'évaluation**

Dans le domaine de la normalisation de l'évaluation, il y a trois niveaux qui se complètent. Nous avons : a) les principes directeurs qui définissent les idéaux poursuivis; b) les normes, règles et standards qui fournissent des prescriptions pour la pratique de l'évaluation; c) l'éthique et la déontologie qui guident la moralité et les comportements professionnels. Dans cette section du chapitre, nous allons d'abord définir des concepts de base. Ensuite, nous procéderons à une revue des normes émises par diverses entités normatives. Enfin, nous mentionnerons les contributions à la normalisation de plusieurs entités d'évaluation.

### **Définition de concepts clés**

Clarifions des concepts clés et définissons des termes utilisés en évaluation.

- a. Les *principes* sont des « règles théoriques générales qui déterminent la conduite et l'action » (LeParisien sensagent, 2016). Ils sont une généralisation acceptée comme vraie et pouvant servir de base à un raisonnement ou à une conduite professionnelle. Ils reflètent les idéaux poursuivis. Ils sont également appelés principes directeurs.



- b. Les *normes* sont des préceptes régissant la pratique de l'évaluation permettant de fournir une garantie de performance et de qualité. Selon l'Organisation internationale pour la standardisation (ISO, s. d.) : « *les normes reposent sur les connaissances des experts dans leur domaine de prédilection.* » Elles peuvent être établies théoriquement ou empiriquement conformément à une moyenne. Elles sont des repères précisant les prescriptions techniques à respecter pour la conduite d'une évaluation. En évaluation, il s'agit d'un des apports évidents de la démarche positiviste, qui est de permettre la conceptualisation de normes. Une norme est un ensemble de règles de conformité ou de fonctionnement qui peut être légiféré par un organisme de normalisation telle l'ISO. Les normes appuient la professionnalisation de l'évaluation en guidant la pratique assurant une qualité et une crédibilité.
- c. Les *règles* sont des « prescriptions, de l'ordre de la pensée ou de l'action, qui s'imposent dans un cas donné » (Larousse, s. d.). En évaluation, il s'agit de consignes utiles qui permettent une qualité dans l'exercice professionnel de la fonction. Les règles de l'art constituent un « ensemble des règles, écrites ou non écrites, reconnues comme cadre de référence par les experts d'un domaine donné et dont le respect constitue une obligation implicite de l'exercice d'une profession » (OQLF, 2021).
- d. Un *standard* est terme anglais qui correspond à un niveau de qualité (*a level of quality*) (Cambridge Dictionary, s. d.) En évaluation, un standard est un ensemble de recommandations ou de préférences préconisées par un groupe d'utilisateurs avisés, lequel sert à jauger un produit d'évaluation.
- e. Les *critères* sont les « éléments de référence qui permettent de juger, d'estimer, de définir quelque chose » (Larousse, s.d.). Ils peuvent être subjectifs ou objectifs. Souvent, en évaluation, une approche multicritère est utilisée et plusieurs indices servent à la sommation d'une conclusion.
- f. Les *spécifications* font référence aux exigences spécifiques auxquelles doit répondre un produit, un processus ou un service d'évaluation. Les spécifications sont normalement incluses dans les termes de référence d'une évaluation.
- g. Les *règles de conduite* sont fort utiles pour les organisations qui souhaitent se doter de règles générales claires quant aux comportements souhaités en milieu professionnel. Les règles énoncées dans un code de conduite et d'éthique permettent à tous de bien connaître explicitement les comportements attendus.
- h. Le code d'éthique fait référence au texte qui énonce « les valeurs et les principes à connotation morale ou civique auxquels adhère une organisation, et qui servent de guide à un individu ou à un groupe afin de l'aider à juger de la justesse de ses comportements » (OQLF, 2008).
- i. Le code de déontologie est défini comme « un texte réglementaire énonçant les règles de conduite professionnelle qui régissent l'exercice d'une profession ou d'une fonction et faisant état des devoirs, des obligations et des responsabilités auxquels sont soumis ceux qui l'exercent » (OQLF, 2007).

## Références internationales de la normalisation de l'évaluation

Voici des références séminales quant à la normalisation de l'évaluation.

### Comité mixte sur les normes d'évaluation en éducation

Le comité mixte sur les normes d'évaluation en éducation (CMNEE) (*Joint Committee on Standards for Educational Evaluation*) a produit une première norme intitulée *Standards for Evaluations of Educational Programs, Projects and Materials* (1981). Ce comité a établi une fondation solide au développement des normes professionnelles. La norme fut actualisée par le CMNEE et publiée en 1994 comme *The program evaluation standards : How to assess evaluations of educational programs*. En 2011, en collaboration avec la Société canadienne d'évaluation (SCE) le CMNEE a publié partiellement en français sa troisième version du *The Program Evaluation Standards : A Guide for Evaluators and Evaluation Users*. La SCE utilise les normes du comité mixte comme les siennes en ce qui a trait aux normes d'évaluation de programme.

Les normes du CMNEE comprennent cinq normes principales. Elles sont :

- Les *normes d'utilité* qui visent à aider les parties prenantes des interventions à trouver des processus et des produits d'évaluation correspondant à leurs besoins;
- Les *normes de faisabilité* qui visent à augmenter l'efficacité et le rendement de l'évaluation;
- Les *normes de convenance* qui désignent ce qui est convenable, équitable, légal, correct et juste dans les évaluations;
- Les *normes de précision* qui visent à augmenter la fiabilité et la véracité des représentations, des propositions et des constats d'évaluation, surtout s'agissant de ceux et celles qui contribuent aux interprétations et aux jugements portant sur la qualité;
- Les *normes d'imputabilité* qui encouragent la documentation adéquate à l'appui des évaluations et une perspective méta-évaluative axée sur l'amélioration et l'imputabilité des processus et produits d'évaluation.

Les critères des cinq normes sont récapitulés dans les tableaux qui suivent<sup>1</sup>. Ils sont une excellente référence pour les praticiens d'évaluation.

---

1. Les auteurs ont dûment informé la Société canadienne d'évaluation et le CMNEE de la reproduction de la nomenclature et des énoncés de ces normes dans ce chapitre.

**Tableau 1 — Normes d'évaluation de programmes du Comité mixte sur les normes d'évaluation en éducation adoptées par la Société canadienne d'évaluation en janvier 2012 (Yarbrough *et al.*, 2011, traduit par la Société canadienne d'évaluation, 2015).**

Normes d'utilité	
U1 Crédibilité de l'évaluateur ou de l'évaluatrice	Les évaluations devraient être menées par des personnes compétentes qui instaurent et maintiennent leur crédibilité dans le contexte d'évaluation.
U2 Considération des parties prenantes	Les évaluations devraient accorder de l'attention à tous les individus et groupes qui sont investis dans l'intervention et qui sont affectés par son évaluation.
U3 Objectifs négociés	Les objectifs de l'évaluation devraient être identifiés et continuellement négociés selon les besoins des intéressés.
U4 Valeurs explicites	Les évaluations devraient clarifier et spécifier les valeurs individuelles et culturelles qui sous-tendent les finalités, les processus et les jugements évaluatifs.
U5 Informations pertinentes	Les informations d'évaluation devraient répondre aux besoins connus et émergents des intéressés.
U6 Processus et produits importants	Les évaluations devraient mettre en œuvre des activités et formuler des descriptions et des jugements encourageant les participants à redécouvrir, réinterpréter ou revoir leur compréhension et leurs comportements.
U7 Communication et rapports opportuns et appropriés	Les évaluations devraient tenir compte des besoins d'information persistants de leurs multiples publics.
U8 Préoccupation des conséquences et de l'influence	Les évaluations devraient promouvoir leur utilisation responsable et adaptée, tout en évitant les retombées négatives et les usages fautifs involontaires.
Normes de faisabilité	
F1 Gestion de projet	Les évaluations devraient s'appuyer sur des stratégies efficaces de gestion de projet.
F2 Procédures pratiques	Les procédures d'évaluation devraient être pratiques et adaptées au fonctionnement de l'intervention.
F3 Viabilité contextuelle	Les évaluations devraient prendre acte des intérêts et besoins politiques et culturels des individus et des groupes, et devraient les suivre et en assurer une représentation équilibrée.
F4 Utilisation des ressources	Les évaluations devraient être rentables et efficaces dans leur utilisation des ressources.
Normes de convenance	
C1 Orientation adaptée et universelle	Les évaluations devraient être sensibles aux intervenants et à leurs communautés.
C2 Accords officiels	Les accords officiels devraient être négociés afin de clarifier les obligations des clients et d'autres intéressés, et tenir compte de leurs besoins, attentes et contextes culturels.
C3 Droits de la personne et respect	Les évaluations doivent être élaborées et menées afin de protéger les droits de la personne et légaux, et afin de préserver la dignité des participants et des divers intéressés.
C4 Clarté et justesse	Les évaluations devraient être intelligibles et justes par rapport aux besoins et aux objectifs des parties prenantes.
C5 Transparence et divulgation	Les évaluations devraient offrir à tous les intéressés des descriptions détaillées des constats, des limites et des conclusions, à moins que cette procédure ne contrevienne aux obligations légales et de convenance.
C6 Conflits d'intérêts	Les évaluations devraient permettre d'identifier et d'aborder, ouvertement et honnêtement, les conflits d'intérêts réels ou perçus qui risquent de les compromettre.
C7 Responsabilité financière	Les évaluations devraient rendre compte de toutes les ressources utilisées et respecter toutes les procédures et tous les processus financiers robustes.

Normes de précision	
P1 Conclusions et décisions justifiées	Les conclusions et les décisions relatives à l'évaluation devraient être ouvertement justifiées en regard des cultures et des contextes où elles sont importantes.
P2 Information valide	L'information servant les fins de l'évaluation devrait respecter les objectifs de départ et étayer la validité des interprétations.
P3 Information fiable	Les procédures de l'évaluation devraient produire des informations suffisamment fiables et cohérentes aux fins de l'utilisation qui en est prévue.
P4 Descriptions claires de l'intervention et du contexte	Les interventions et leurs contextes devraient être décrits de manière assez détaillée et exhaustive pour les fins de l'évaluation.
P5 Gestion de l'information	Les évaluations devraient faire appel à des méthodes systématiques de collecte, de révision, de vérification et de stockage des informations.
P6 Analyses et concepts robustes	Les évaluations devraient s'appuyer sur des concepts et des analyses techniquement adéquats et appropriés pour les fins de l'évaluation.
P7 Raisonnement clair de l'évaluation	Le raisonnement d'évaluation qui est suivi à partir de l'information et de son analyse et qui aboutit à des constats, des interprétations, des conclusions et des jugements devrait être documenté de manière claire et exhaustive.
P8 Communication et rapports	Les communications relatives à l'évaluation devraient avoir une portée appropriée et comporter des mesures pour les sauvegarder des conceptions erronées, des biais, des distorsions et des erreurs.
Normes d'imputabilité en évaluation	
I1 Documentation relative à l'évaluation	Les évaluations devraient documenter de façon exhaustive leurs objectifs négociés, la façon dont leurs schémas et procédures ont été appliqués, et leurs données et résultats.
I2 Méta-évaluation interne	Les évaluateurs et évaluatrices devraient utiliser ces normes, et d'autres normes applicables, pour juger du caractère responsable du schéma d'évaluation, des procédures utilisées, de l'information collectée et des résultats de l'évaluation.
I3 Méta-évaluation externe	Les commanditaires, les clients, les évaluateurs et évaluatrices ainsi que les autres actrices et acteurs intéressés par l'évaluation de l'intervention devraient encourager l'application de méta-évaluations externes en utilisant ces normes, et d'autres normes applicables.

## Organisation pour la coopération et le développement économique – Comité d'aide au développement

Le Comité d'aide au développement de l'OCDE a émis en 1991 les *Principes pour l'évaluation de l'aide au développement*. Cinq critères d'évaluation ont été retenus et sont devenus étalons de l'évaluation à l'international. Ils s'agit de : la pertinence, l'efficacité, l'efficience, l'impact, la viabilité/durabilité. Le Réseau du CAD sur l'évaluation du développement (EvalNet), en tant qu'organe subsidiaire du CAD, chargé d'apporter un appui à l'évaluation et au renforcement de l'apprentissage et l'élaboration des politiques, a effectué un travail de révision, paru en 2019, à la suite de 25 ans d'utilisation de ces critères et dans le cadre de la mise en œuvre de l'Agenda 2030 et de l'Accord de Paris dans le cadre de la Convention-cadre des Nations Unies sur les changements climatiques. Cette révision faisait suite aux critiques des parties prenantes de la coopération internationale, notamment sur l'orientation des critères originaux sur des dimensions « projets » et la non-prise en compte suffisante des questions telles que la complexité et les arbitrages entre priorités, l'équité ou encore l'intégration des droits de la personne et l'égalité des sexes. Plusieurs évaluateurs et évaluatrices de la coopération internationale considèrent que ces critères demeurent encore faibles pour prendre en compte les enjeux complexes du développement durable, notamment en termes de réflexion sur la pertinence, l'ampleur du changement et l'aspect systémique des changements à porter. Ces critères demeurent très attachés à la notion d'intervention sans dépasser l'évaluation des changements (en lien avec la résolution des problèmes complexes) « sans intervention » (Patton, 2020).

EvalNet a défini six critères d'évaluation et deux principes d'utilisation. Les critères sont destinés à guider les évaluations, et ont été affinés en vue d'améliorer la qualité et l'utilité des évaluations et de renforcer la contribution de l'évaluation au développement durable. Les critères sont :

- La pertinence (L'intervention répond-elle au problème?)
- La cohérence (L'intervention s'accorde-t-elle avec les autres interventions menées?)
- L'efficacité (L'intervention atteint-elle ses objectifs?)
- L'efficience (Les ressources sont-elles utilisées de manière optimale?)
- L'impact (Quelle différence l'intervention fait-elle?)
- La viabilité/durabilité (Les bénéfices sont-ils durables?)

Voici les deux principes fondamentaux définis pour guider la manière dont les critères sont utilisés :

- « Les critères devraient être appliqués de façon réfléchie afin de promouvoir la qualité et l'utilité des évaluations » (CAD, 2019, p. 6); ce qui renvoie à la prise en compte du contexte, notamment du cadre spécifique de l'évaluation, du type d'intervention évaluée et des parties prenantes.
- « L'utilisation des critères dépend de la finalité de l'évaluation » (CAD, 2019, p. 6); ce qui renvoie à la nécessité d'adapter les critères aux besoins des parties prenantes et au contexte de l'évaluation.

Il est important que les critères soient appliqués dans un contexte large et en lien avec d'autres principes et lignes directrices visant à éclairer la contribution des parties prenantes sur la façon de réaliser des évaluations utiles et de qualité.

En 2002, EvalNet a publié le *Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats* en anglais, en espagnol et en français. C'est un document clé pour l'harmonisation des termes d'évaluation à l'international. Il demeure encore, à ce jour, un document de référence.

En 2010, le CAD de l'OCDE a émis les *Normes de qualité pour l'évaluation du développement*, qui ont pour but « d'améliorer la qualité des processus d'évaluation du développement et des produits qui en découlent ainsi que faciliter la collaboration. Elles détaillent les éléments clés qui font la qualité de chacune des phases d'une démarche type d'évaluation, à savoir la définition de la finalité de l'évaluation, la planification et la conception de cette dernière, sa mise en œuvre et la présentation de ses résultats, et enfin l'examen des leçons à tirer et l'utilisation des résultats » (CAD, 2010ab).

### Le Groupe des Nations Unies pour l'évaluation

Les Normes et règles d'évaluation du Groupe des Nations Unies pour l'évaluation (UNEG), adoptées en 2005, puis actualisées et disséminées en juin 2016, ont été conçues pour être appliquées pour l'ensemble des évaluations du système des Nations Unies afin de clarifier les procédures dans le contexte du Programme de développement durable à l'horizon 2030. Par leur positionnement, elles ont permis de renforcer et d'harmoniser les pratiques d'évaluation comme une référence majeure dans la communauté des évaluateurs et évaluatrices. Les dix normes générales de l'UNEG sont les suivantes :

Norme 1 – Principes, objectifs et cibles	Norme 2 – Utilité
Norme 3 – Crédibilité	Norme 4 – Indépendance
Norme 5 – Impartialité	Norme 6 – Éthique
Norme 7 – Transparence	Norme 8 – Droits de l'homme et égalité des sexes
Norme 9 – Capacités d'évaluation nationales	Norme 10 – Professionnalisme

Le Groupe des Nations Unies pour l'évaluation a œuvré intensément vers une professionnalisation de la fonction d'évaluation au sein des Nations Unies. Voici une sélection de documents de référence produits par l'UNEG :

- 2008 – Compétences de base pour les évaluateurs des Nations Unies;
- 2008 – Code de conduite éthique – révisé en 2020;
- 2010 – Contrôle de la qualité des mandats et rapports initiaux d'évaluation;
- 2010 – Contrôle de la qualité des rapports d'évaluation;
- 2010 – Bonnes pratiques pour le suivi aux évaluations;
- 2012 – Évaluer la coopération pour le développement – Récapitulatif des normes et standards de référence;
- 2013 – Manuel de l'UNEG pour la conduite de l'évaluation du travail normatif dans le système des Nations Unies;
- 2014 – Intégration des droits humains et de l'égalité des sexes dans les évaluations;
- 2016 – Cadre de compétences d'évaluation;

- 2018 – Guide pour l'évaluation de l'intégration du genre au niveau institutionnel;
- 2019 – Des meilleurs critères pour des meilleures évaluations – Définitions adaptées et principes d'utilisation.

### Associations volontaires de professionnels en évaluation

Il y a une contribution significative faite par les associations d'évaluation aux niveaux national, régional et mondial encourageant la normalisation, l'institutionnalisation et la professionnalisation qu'on ne saurait synthétiser dans le présent chapitre. Ceci est reflété dans d'autres chapitres de l'ouvrage. Pour avoir un aperçu de ce qui se fait, le lecteur est invité à visiter le site Internet de l'Organisation internationale de coopération en évaluation (OICE) et celui de *EvalPartners*.

Toutefois, il y aurait un manquement important à ne pas mentionner ici les contributions du Réseau francophone de l'évaluation. Le RFE est une tête de réseau regroupant des associations d'une vingtaine de pays lancée en 2008 à Québec. Ses objectifs sont : a) développer l'offre francophone d'évaluation; b) constituer un corpus théorique et technique en langue française; c) développer entre les associations nationales d'évaluation une coopération active; d) promouvoir l'utilisation des résultats de l'évaluation dans le processus de décision publique.

De plus, il existe plusieurs associations et réseaux régionaux d'évaluation. Nous allons seulement les énumérer, à savoir :

- La Communauté des évaluateurs de l'Asie du Sud (CoE);
- La Société australasienne d'évaluation (AES);
- La Société européenne d'évaluation (SEE);
- Le Réseau d'évaluation de l'Afrique du Nord et du Moyen-Orient (EvalMENA);
- Le Réseau international d'évaluation de programme (IPEN) de l'Europe de l'Est;
- Le Réseau régional de suivi et évaluation de l'Amérique latine et des Caraïbes (ReLAC).

Enfin, il y a lieu de mentionner un réseau mondial anglophone d'apprentissages et d'évaluations dans le domaine des interventions humanitaires. Il s'agit de l'ALNAP (*Active Learning Network for Accountability and Performance*) qui a développé des principes, des normes, des standards et une déontologie. L'ALNAP facilite la collaboration entre une trentaine d'organisations mondiales. Le site Internet de l'ALNAP offre 21 100 ressources aisément accessibles.

Il y a sans doute plusieurs associations et réseaux nationaux d'évaluation qui sont proactifs pour la promotion et le développement de la normalisation et de la professionnalisation de l'évaluation. Mentionnons l'Association américaine d'évaluation (AEA), l'Association marocaine d'évaluation (AME), laquelle a contribué à l'incorporation en 2011, dans la nouvelle constitution du pays, à l'obligation d'évaluer les investissements publics. Méritent aussi d'être mentionnés : le Réseau nigérien de suivi et évaluation (ReNSE) et l'Association sénégalaise d'évaluation (SenEVAL).

Compte tenu des limitations d'espace du présent chapitre, nous allons nous pencher sur trois associations d'évaluation qui ont contribué de façon singulière et systématique au domaine de la normalisation de l'évaluation. Il s'agit de l'AfrEA, IDEAS et la SCE.



## **L'AfrEA**

L'Association africaine d'évaluation est une organisation faitière de plus de 45 associations et réseaux nationaux de suivi et d'évaluation en Afrique. Le leitmotiv de l'AfrEA est de « *promouvoir l'évaluation conçue et pilotée par l'Afrique* ». Le concept vise à identifier et à développer une approche à l'évaluation qui soit singulièrement africaine. L'accent est mis sur le contexte, la culture, l'histoire et les croyances qui façonnent la nature des évaluations, particulièrement dans une réalité africaine qui est fort diversifiée et complexe. Par conséquent, l'AfrEA, créée en 1999, veut formuler des principes d'évaluation qui outrepassent la technicité des normes pour inclure une préoccupation humaine qui reflète les dimensions socioculturelles africaines.

L'Association africaine d'évaluation a émis en avril 2020 une ébauche des Principes directeurs de l'évaluation en Afrique aux fins de consultation et de discussion. Ces principes « visent à favoriser une évaluation qui soit : réellement utile et responsabilisante pour les parties prenantes et les populations en Afrique—grâce à la redevabilité, à de nouvelles perspectives et à l'appropriation de l'évaluation et des changements y afférents; techniquement robuste et efficace; éthiquement correcte; centrée sur l'Afrique, mais ouverte sur le monde; connectée au monde, et donc soucieuse des préoccupations du monde; centrée sur les opportunités et les solutions adaptées à cette époque, où l'empreinte de l'humanité sur le monde appelle de nouvelles idées et connaissances pour un changement transformateur, du niveau local au niveau global » (AfrEA, 2020, p. 6).

Cinq principes clés fournissent un cadre pour les pratiques d'évaluation menées en Afrique et pour l'Afrique. Vingt-deux principes de mise en œuvre élaborent les cinq principes clés. Le tableau qui suit résume ces principes :

**Tableau 2 — Les principes directeurs de l'évaluation en Afrique.**  
Reproduit de AfrEA, 2020, p. 6-7.

P. Puissante pour les Africains	P1. Mener un processus approprié et habilitant; P2. Encourager la réciprocité, y compris la responsabilité mutuelle; P3. Susciter l'apprentissage pour acquérir des idées utiles; P4. Valoriser et renforcer les capacités locales.
T. Techniquement robuste	T1. Être systématique et analytique; T2. Être transparente et claire; T3. Connaître les (pré) dispositions; T4. Assurer une évaluation réalisable; T5. Être efficiente; T6. Être culturellement sensible.
E. Éthiquement saine	E1. Être sensible aux parties prenantes et aux relations; E2. Protéger les droits des personnes; E3. Garantir la diversité et l'inclusion; E4. Remédier aux inégalités et aux asymétries de pouvoir; E5. Se départir des intérêts acquis; E6. Considérer les compromis.



A. Centrée sur l'Afrique, mais ouverte sur le monde	A1. S'engager sur des questions importantes en Afrique; A2. Considérer les cadrages et les méthodes de l'Afrique; A3. Apprendre des pays du Sud, des communautés autochtones, et d'autres contextes.
C. Connectée au monde	C1. Reconnaître l'interdépendance et l'interconnexion; C2. Favoriser l'évaluation de la durabilité en conformité aux principaux accords mondiaux et à la nécessité de sauvegarder la nature; C3. S'efforcer de contribuer au besoin urgent d'un changement durable et transformateur.

## IDEAS

Créée en 2002, l'Association d'évaluation du développement international est une association mondiale de 700 membres individuels venant de plus de 90 pays. C'est une association qui fonctionne présentement essentiellement en anglais, d'où son sigle IDEAS pour *International Development Evaluation Association*. IDEAS focalise ses efforts sur l'évaluation internationale du développement durable. L'association est très engagée à renforcer la profession d'évaluation, à appuyer le développement des capacités, à promouvoir l'avancement des théories, des méthodes et de la pratique d'évaluation et son utilisation. Voici les principaux travaux appuyant la normalisation de l'évaluation internationale :

- 2012 – Compétences pour les évaluateurs d'évaluation du développement, pour les gestionnaires et les mandataires;
- 2014 – Code d'éthique;
- 2014 – Position proposée par IDEAS concernant les qualifications en évaluation et rôle possible pour IDEAS;
- 2019 – Déclaration de Prague qui inclut l'appui à la professionnalisation et le développement des capacités;
- 2019 – Considération de l'établissement d'une Académie internationale d'évaluation;
- 2021 – L'évaluation transformatrice.

## La SCE

La Société canadienne d'évaluation est très dynamique dans la normalisation et la professionnalisation de l'évaluation. En 2021, elle comptait 1 700 membres. Plus de 500 sont des évaluatrices et évaluateurs qualifiés. Elle propose un premier référentiel de compétences professionnelles pour l'exercice de l'évaluation en 2010, puis le met à jour en 2018 (SCE, 2018). Les compétences y sont définies comme « le fondement, les connaissances, les compétences et les attitudes que les évaluateurs de programme doivent posséder pour produire des évaluations solides » (Stevahn *et al.*, 2005, traduit librement par la Société québécoise d'évaluation de programme (SQEP, 2022)). Il y a 36 compétences réparties en cinq domaines :

- Celles d'**une pratique réflexive** (en lien avec les connaissances théoriques et pratiques, le recours aux normes et aux lignes de conduite et à l'autoréflexion sur sa pratique professionnelle, incluant le renforcement des capacités pour ses compétences);
- Celles d'**une pratique technique** (renvoie aux décisions stratégiques, méthodologiques et interprétatives);

- Celles d'**une pratique contextuelle** (lié à l'adaptation et au contexte de l'évaluation qui la rendent unique);
- Celles d'**une pratique de gestion** (en lien avec les aptitudes en gestion d'un projet d'évaluation);
- Celles d'**une pratique concernant les relations interpersonnelles** (en lien avec les habiletés sociales et personnelles requises pour communiquer et interagir efficacement avec les parties prenantes).

Le référentiel complet est disponible sur le site de la SCE : <https://evaluationcanada.ca/fr/competences-pour-les-evaluateurs-canadiens>

La SCE propose également un guide d'évaluation pour une éthique de l'évaluation réalisé à l'issue de deux rondes de commentaires recueillis auprès des membres d'un groupe d'experts en éthique de l'évaluation. Ce guide fait le lien entre le référentiel des compétences de l'évaluateur ou de l'évaluatrice, les normes d'évaluation issues du Comité mixte sur les normes d'évaluation en éducation, présentés dans les tableaux ci-dessus, et les valeurs fondamentales promues par la SCE et qui proviennent de la certitude que les «évaluateurs valorisent par-dessus tout le travail bien fait» selon l'expression de Love, 2018 :

- «Les droits et le bien-être des personnes et des peuples»;
- «La recherche de la vérité, la franchise et la transparence»;
- «La responsabilité envers les parties prenantes et la société». (SCE, 2020, p. 6).

Ce guide est également disponible sur le site de la SCE : <https://evaluationcanada.ca/fr/ethique>.

L'ensemble de ces références permet à la SCE de soutenir les efforts de professionnalisation de la pratique de l'évaluation au Canada, notamment à travers la conception du titre d'évaluateur qualifié (ÉQ), qui définit, reconnaît et fait la promotion «d'un exercice professionnel qui répond aux plus hauts standards d'éthique, de qualité et de compétence au Canada» (SCE, 2014). Le Programme des titres professionnels est composé de trois piliers : un code de déontologie, des normes d'évaluation et un ensemble de compétences. À l'issue d'un processus rigoureux, exigeant et introspectif, le détenteur du titre «a fourni une preuve convaincante qu'il possède les aptitudes, les connaissances et l'expérience professionnelle requises identifiées par la SCE comme celles nécessaires pour être considéré comme un évaluateur compétent» (SCE, 2014). Il est important de souligner que ce terme est un titre de compétences, et non un certificat ou un brevet. De plus amples informations sont disponibles sur le site Web de la SCE : <https://evaluationcanada.ca/fr/propos-du-titre-deq>

### **Contribution de sociétés privées d'évaluation et d'évaluateurs et évaluatrices autonomes**

Nous voulons signaler qu'ils sont fort nombreux les chercheurs, académiciens et praticiens qui ont contribué et innové en faveur de la normalisation et de la professionnalisation de l'évaluation. Ces personnes viennent de l'ensemble des disciplines académiques et professionnelles. Certains gurus dominent l'évolution de la pensée en évaluation.

Il y a également plusieurs entités de recherche, comme le Centre de recherche en développement international, qui appuient la réflexion. Des sociétés de consultants en évaluation ont rendu disponibles gratuitement via Internet des outils utiles pour les praticiens en évaluation. Un bel exemple est *BetterEvaluation*, qui a fourni une plateforme de connaissances et d'échange, un soutien pour le développement organisationnel en évaluation et la promotion et le suivi de projets de recherche pour le développement de l'évaluation.

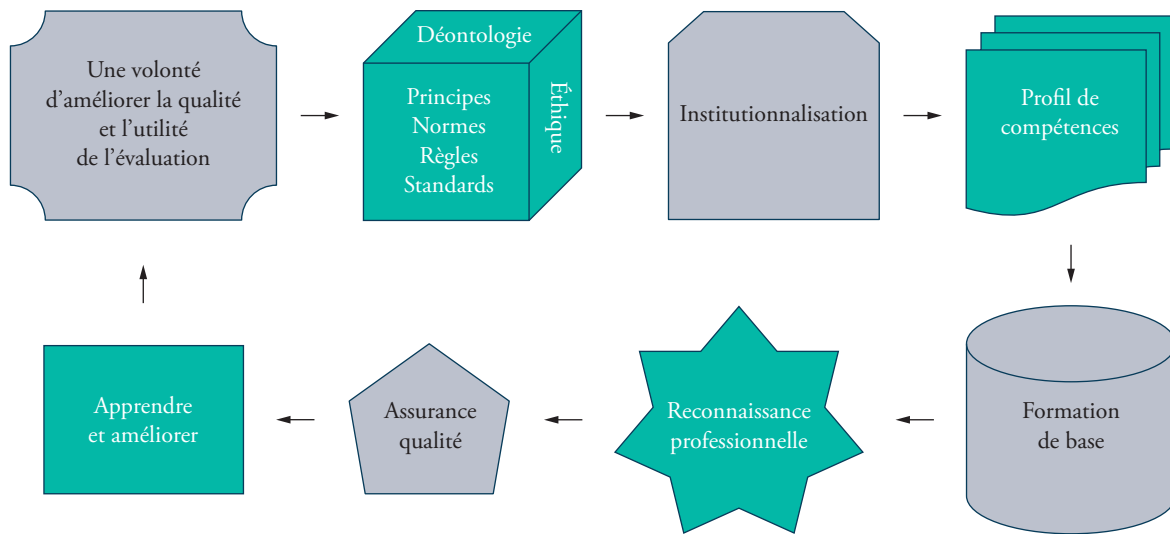
## Utilité de la normalisation de l'évaluation

Les principes, normes, règles et standards, codes d'éthique et déontologie ont tous comme finalité la production d'évaluations qui sont le résultat d'une approche rigoureuse basée sur l'évidence et qui reflète un comportement professionnel, intègre et équitable générant crédibilité et confiance.

La professionnalisation repose sur la normalisation des pratiques évaluatives. S'il y a normalisation et professionnalisation, les praticiens et praticiennes d'évaluation pourront recevoir une reconnaissance comme personne qualifiée à conduire une évaluation. Il s'ensuit que le respect des principes et des normes d'évaluation renforce la rigueur méthodologique qui habilite des travaux d'évaluation de qualité. Le respect des principes et des normes est essentiel pour une crédibilité et l'utilisation des résultats d'évaluation aux fins de prise de décision et d'apprentissage institutionnel.

Un professionnalisme de l'évaluation nécessite non seulement le respect des principes, des normes, des règles et des standards, mais aussi un comportement qui respecte le code d'éthique et la déontologie. Il faut qu'il y ait également une institutionnalisation de l'évaluation. Par exemple, un pays peut inclure dans sa constitution nationale l'obligation d'évaluer. Un gouvernement peut décréter une politique nationale d'évaluation qui donne aux entités publiques des consignes d'évaluation. Une organisation peut se doter d'une politique d'évaluation qui régit la fonction. Dans la pratique, le professionnalisme se traduit par des profils de compétences, lesquels peuvent être incorporés dans des descriptions de poste utiles au moment du recrutement. Le professionnalisme repose sur un savoir professionnel qui est acquis grâce à une formation spécialisée. Il devient alors facile d'octroyer une reconnaissance professionnelle aux évaluateurs et évaluatrices. Les normes de qualité sont pertinentes et utiles au moment d'assurer la qualité et l'utilité des évaluations. Il va de soi que si les évaluations permettent d'apprendre des interventions et de les améliorer, il en ressortira une volonté accrue d'assurer la qualité et la valeur ajoutée des évaluations subséquentes. Somme toute, le respect des usagers d'évaluation donnera à l'évaluation une indépendance et un meilleur positionnement.

Ci-après, un schéma représente la place importante que joue la normalisation dans un modèle logique de la professionnalisation.



**Figure 1**

Contribution des normes à la professionnalisation.

## Conclusion

Une évaluation de qualité nécessite une démarche évaluative qui soit rigoureuse, crédible et fiable. Pour ce faire, il faut des normes d'évaluation, un code de déontologie, un ensemble de compétences professionnelles et une conduite professionnelle qui répond aux plus hauts standards d'éthique. Chacune de ces variables ne suffit pas en elle-même, il faut une conjugaison de toutes ces variables. De surcroît, il doit y avoir un souci d'autoévaluation pour une amélioration professionnelle constante.

La revue des différentes initiatives de normalisation nous amène à conclure qu'un travail énorme a été accompli. Au début, les principes, normes et règles étaient empruntés au domaine de la recherche scientifique. C'est ce qui leur a donné rigueur et crédibilité. Les contributions de l'AfrEA ont démontré que la qualité et l'utilité des évaluations ne reposent pas uniquement sur les approches méthodologiques, mais beaucoup plus sur les facteurs humains qui imprègnent la réflexion évaluative. Les nombreuses approches à la normalisation révèlent que ce domaine est en devenir et qu'il s'aligne vers une plus grande harmonisation qui sera le résultat de la poursuite d'une réflexion plus systémique et systématique sur le plan international. Ainsi, des entités comme le système des Nations Unies, l'Organisation internationale de coopération en évaluation et l'Association d'évaluation du développement international sont appelées à assumer un rôle de leadership. En terminant, signalons la contribution importante que peuvent jouer les évaluatrices et évaluateurs émergents et les nombreuses communautés de pratique à la normalisation des nouvelles tendances en évaluation.

## Exercice de réflexion

Déterminer les principes, les normes, les règles, les standards, l'éthique et la déontologie d'évaluation.

1. Est-ce que les droits humains font partie des normes d'évaluation ?
2. Quelle est la différence entre une norme et une règle d'évaluation ?
3. Quels critères d'évaluation doivent être inclus au minimum dans les termes de référence ?
4. Existe-t-il une norme d'évaluation pour l'égalité des sexes ?
5. Quelle est la valeur fondamentale de la déontologie en évaluation ?
6. Comment la culture influence-t-elle l'application des normes d'évaluation ?
7. Devrait-on exiger une professionnalisation des personnes exerçant l'évaluation ?

## Références bibliographiques

- Association africaine d'évaluation (AfrEA). (2020). *Principes directeurs pour l'évaluation en Afrique version 2020*. [https://afrea.org/aeg/AEG\\_FRENCH.pdf](https://afrea.org/aeg/AEG_FRENCH.pdf)
- CAD. Comité d'aide au développement. (1991). *Principes pour l'évaluation de l'aide au développement*. OCDE. [http://www.eval.fr/wp-content/uploads/2018/02/EVAL\\_CAD-OCDE\\_principes-1991.pdf](http://www.eval.fr/wp-content/uploads/2018/02/EVAL_CAD-OCDE_principes-1991.pdf)
- CAD. Comité d'aide au développement. (2002). *Glossaire des principaux termes relatifs à l'évaluation et à la gestion axée sur les résultats*. OCDE. <https://www.oecd.org/development/evaluation/2754804.pdf>
- CAD. Comité d'aide au développement. (2010a). *Évaluer la coopération pour le développement. Récapitulatif des normes et standards de référence (2<sup>e</sup> éd.)*. OCDE.
- CAD. Comité d'aide au développement. (2010b). *Normes de qualité pour l'évaluation du développement*. OCDE. <https://www.oecd.org/fr/cad/evaluation/normesdequalitepourledeveloppement.htm#:~:text=Les%20Normes%20de%20qualit%C3%A9%20pour,que%20de%20faciliter%20la%20collaboration>
- CAD. Comité d'aide au développement. (2019). *Des meilleurs critères pour des meilleures évaluations – Définitions adaptées et principes d'utilisation*. OCDE. <http://www.oecd.org/fr/cad/evaluation/criteres-adaptees-evaluation-dec-2019.pdf>
- Cambridge Dictionary. (s. d.) Standard. Dans *Cambridge Dictionary*. <https://dictionary.cambridge.org/fr/dictionnaire/anglais/standard>
- IDEAS. (2012). *Competencies for Development Evaluation Evaluators, Managers and Commissioners*. [https://drive.google.com/file/d/1pjRq4M5\\_76hzntDcm9zO\\_Fq053jwODE8/view](https://drive.google.com/file/d/1pjRq4M5_76hzntDcm9zO_Fq053jwODE8/view)
- IDEAS. (2014a). *Code d'éthique*. <https://drive.google.com/file/d/0BzggS7JkIKT6M1BJTUhLd0lPTGc/view>
- IDEAS. (2014b). *Position proposée pour IDEAS concernant les qualifications en évaluation et rôle possible pour IDEAS – Proposed Position on Qualification in Evaluation and Possible Role for IDEAS*. <https://drive.google.com/drive/folders/0B9uTDrpkqkkqZkV5NVBKY2NJaEk>

- IDEAS. (2019). *Report from the IDEAS Survey on an International Evaluation Academy*. <https://drive.google.com/file/d/1vGl94J8CrRmyxTWvrVBcE2JYF0f9Zzhz/view>
- IDEAS et al. (2019). *Prague Declaration on Evaluation for Transformational Change*. <https://drive.google.com/file/d/145-9FBQ07fe1zjqVTgwYs19-IuhtRq8y/view>
- Joint Committee on Standards for Educational Evaluation. (1981). *Standards for Evaluations of Educational Programs, Projects, and Materials*. McGraw-Hill.
- Larousse. (s. d.) Critère. Dans *Dictionnaire en ligne*. <https://www.larousse.fr/dictionnaires/francais/crit%C3%A8re/20567>
- Larousse. (s. d.). Règle. Dans *Dictionnaire en ligne*. <https://www.larousse.fr/dictionnaires/francais/r%C3%A8gle/67653>
- LeParisien sensagent. (2016). *Principes*. Dans *LeParisien sensagent*. <http://dictionnaire.sensagent.leparisien.fr/Principes/fr-fr/>
- Office québécois de la langue français (OQLF). (2021). *Règles de l'art*. Le grand dictionnaire terminologique. [http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id\\_Fiche=8362238](http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8362238)
- Office québécois de la langue français (OQLF). (2008). *Code d'éthique*. Le grand dictionnaire terminologique. [http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id\\_Fiche=8365027](http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8365027)
- Office québécois de la langue français (OQLF). (2007). *Code de déontologie*. Le grand dictionnaire terminologique. [http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id\\_Fiche=8364446](http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8364446)
- OICE. Organisation internationale de coopération en évaluation. (2021). *Inventaire des VOPE, VOPE Directory*. <https://ioce.net/vope-directory>
- Organisation internationale pour la standardisation (ISO). (s. d.). *Normes*. <https://www.iso.org/fr/standards.html>
- Patton, M. P. (2020, 20 mars). *Evaluation Criteria for Transformation : Webinar Summary*. Blue Marble Evaluation. <https://bluemarbleeval.org/latest/evaluation-criteria-transformation-webinar-summary>
- Société canadienne d'évaluation (SCE). (2015). *Normes d'évaluation de programmes*. <https://evaluationcanada.ca/fr/normes-devaluation-de-programmes>
- Société canadienne d'évaluation (SCE). (2014). *Éthique*. <https://evaluationcanada.ca/fr/ethique>
- Société canadienne d'évaluation (SCE). (2014). *Compétences pour les évaluateurs canadiens*. <https://evaluationcanada.ca/fr/competences-pour-les-evaluateurs-canadiens>
- Société canadienne d'évaluation (SCE). (2014). *À propos du titre d'ÉQ*. <https://evaluationcanada.ca/fr/propos-du-titre-deq>
- Société canadienne d'évaluation (SCE). (2018). *Référentiel des compétences professionnelles requises à l'exercice d'évaluation de programmes au Canada*. [https://evaluationcanada.ca/txt/2\\_competences\\_cdn\\_pratique\\_evaluation.pdf](https://evaluationcanada.ca/txt/2_competences_cdn_pratique_evaluation.pdf)
- Société canadienne d'évaluation (SCE). (2020). *Guide de la Société canadienne d'évaluation pour une éthique de l'évaluation*. [https://evaluationcanada.ca/sites/default/files/CES\\_Guidance\\_for\\_Ethical\\_Evaluation\\_approved\\_09\\_01\\_20\\_French-final.pdf](https://evaluationcanada.ca/sites/default/files/CES_Guidance_for_Ethical_Evaluation_approved_09_01_20_French-final.pdf)

- Société québécoise d'évaluation de programme (SQEP). (2022). Compétences pour les évaluateurs. [https://www.sqep.ca/la-sqep/competences-pour-les-evaluateurs#:~:text=Les%20comp%C3%A9tences%20sont%20d%C3%A9finies%20comme,%2C%202005\)%20%5B1%5D.](https://www.sqep.ca/la-sqep/competences-pour-les-evaluateurs#:~:text=Les%20comp%C3%A9tences%20sont%20d%C3%A9finies%20comme,%2C%202005)%20%5B1%5D.)
- Stevahn, L., King, J., Ghore, G. et Minnema, J. (2005). Establishing essential competencies for program evaluators. *American Journal of Evaluation*, 26(2), 43-59. <https://doi.org/10.1177%2F1098214004273180>
- UNEG. (2008a). *Core Competencies for Evaluators of the UN System*. <http://www.unevaluation.org/document/detail/1915>
- UNEG. (2008b). *UNEG Code of Conduct for Evaluation in the UN System*. <http://www.unevaluation.org/document/detail/100>
- UNEG. (2010a). *UNEG Good Practice Guidelines for Follow up to Evaluations*. <http://www.uneval.org/document/detail/610>
- UNEG. (2010b). *UNEG Quality Checklist for Evaluation Terms of Reference and Inception Reports*. <http://www.uneval.org/document/detail/608>
- UNEG. (2010c). *UNEG Quality Checklist for Evaluation Reports*. <http://www.uneval.org/document/detail/607>
- UNEG. (2013). *Manuel de l'UNEG pour la conduite de l'évaluation du travail normatif dans le système des Nations Unies*. <http://www.uneval.org/document/detail/1484>
- UNEG. (2016a). *UNEG Evaluation Competency Framework*. <http://www.uneval.org/document/detail/1915>
- UNEG. (2016b). *Norms and Standards for Evaluation*. <http://www.unevaluation.org/2016-Norms-and-Standards>
- UNEG. (2018). *Guidance on Evaluating Institutional Gender Mainstreaming*. <http://www.uneval.org/document/detail/2133>
- UNEG. (2020). *Code de conduite éthique révisé, Ethical Guidelines for Evaluation*. <http://www.uneval.org/document/detail/2866>
- Van der Berg, R. D., Magro, C. et Adrien M.-H. (2021). *Transformational Evaluation for the Global Crises of Our Times*. IDEAS. <https://ideas-global.org/ideas-book-transformational-evaluation-for-the-global-crises-of-our-times-2/>
- Yarbrough, D. B., Shulha, M. L., Hopson, K. R. et Carutgers, A. F. (2011). *The Program Evaluation Standards : A Guide for Evaluators and Evaluation Users* (3<sup>e</sup> édition). SAGE Publications.





# Le suivi et l'évaluation des objectifs de développement durable (ODD) : état des lieux et défis

*Mahmoud Ghouil*

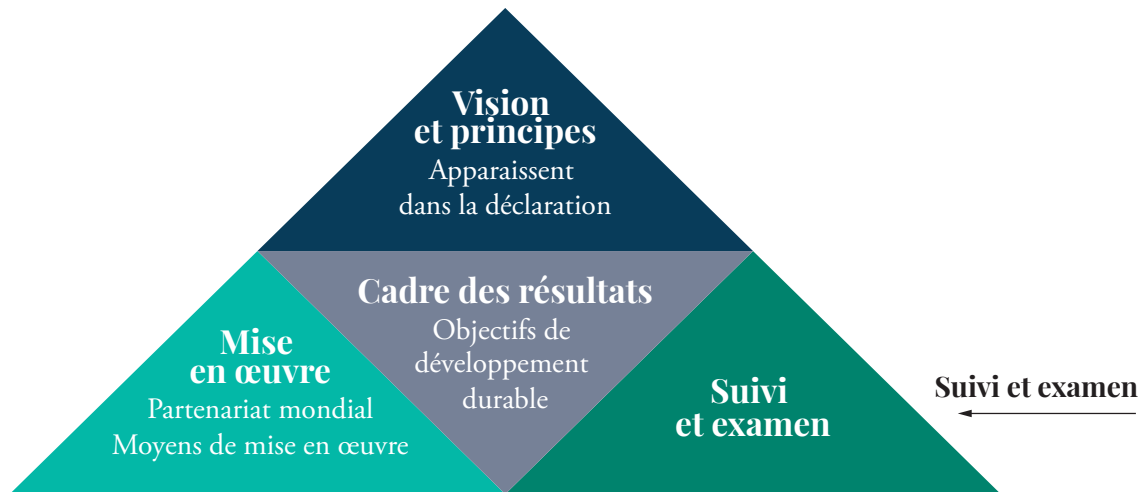
## Introduction

À la différence des objectifs du Millénaire pour le développement (OMD, 2000-2015) qui constituaient un simple cadre des résultats (8 objectifs, 21 cibles) avec des processus d'examen non explicitement liés, l'agenda 2030 pour le développement durable 2016-2030 a prévu dans sa composition non seulement un nouveau cadre des résultats (17 ODD, 169 cibles) avec une vision intégrée, des principes et une stratégie de mise en œuvre, mais surtout un nouveau cadre holistique global et intégré d'examen et de suivi. Cette nouveauté est d'autant plus pertinente qu'elle est susceptible de répondre aux insuffisances constatées des différents mécanismes d'examen non intégrés des OMD. Au cours de ce chapitre, on présentera brièvement un état des lieux du cadre de suivi et d'examen des ODD au niveau mondial pour discuter de la capacité de concilier le double impératif d'adaptation aux besoins spécifiques de chaque pays ainsi qu'aux exigences d'assurer la comparaison internationale entre pays. À la lumière du caractère volontaire des examens au niveau national, on engagera le débat sur la capacité du nouveau cadre du suivi et d'examen à permettre non seulement un suivi régulier des progrès basés sur les indicateurs établis, mais surtout à assurer une fonction évaluative capable de gérer le caractère complexe des ODD. Et ceci, pour expliquer le pourquoi et le comment des changements observés ainsi que leur importance pour les différentes parties prenantes, et notamment pour les personnes laissées pour compte. On se focalisera ainsi sur les préalables à l'évaluation de l'agenda 2030 pour le développement durable dans le monde et au niveau des pays.

## Le cadre du suivi et d'examen des ODD : finalités, principes de mise en œuvre et processus

Fruit d'un engagement ferme de la communauté internationale à assurer un suivi et un examen systématique de la mise en œuvre du programme de développement durable, le nouveau cadre intégré de suivi et d'examen (figure 1) est de nature à apporter une contribution essentielle pour une meilleure réalisation des 17 ODD. Rappelons qu'il permet : 1) une meilleure connaissance et un apprentissage des effets des actions menées sur la population cible; 2) un renforcement des « capacités d'agir » et de pilotage, par une prise de décision basée sur des preuves scientifiques (données probantes); 3) une meilleure transparence et une reddition de comptes des gouvernements. Cette finalité du cadre n'est possible que par la mise en œuvre d'une collecte systématique et continue des données sur des indicateurs spécifiques afin de constituer une base de données fiable. Celle-ci devrait alors permettre de

faire le suivi des progrès réalisés et de favoriser l'identification des éventuels obstacles à l'obtention des résultats. La prise de décision fondée sur des données probantes doit pouvoir favoriser des choix politiques éclairés à condition que les acteurs adaptent leur comportement en fonction des connaissances obtenues. Enfin, les concepts de transparence et de reddition de compte sont plutôt orientés sur la connaissance et l'intérêt des citoyens envers ces actions, ainsi que sur la communication entre les parties prenantes.



**Figure 1**

Les éléments de l'Agenda 2030 pour le développement durable.

Source : UNITAR, 2015, reproduit avec la permission d'UNITAR

Selon la résolution 70/1 de l'Assemblée générale des Nations Unies, la mise en place des mécanismes de suivi et d'examen est guidée par les principes suivants : ils devraient être volontaires et pilotés par les pays (art. 74, paragr. a); ils devraient permettre d'assurer le suivi des progrès accomplis (art. 74, paragr. b) dans une perspective à long terme (art. 74, paragr. c); ils devraient être ouverts, à caractère inclusif, participatif et transparent pour toutes les parties prenantes concernées (art. 74, paragr. d); ces mécanismes devraient être centrés sur la personne, être sensibles à la question de genre, respecter les droits de l'homme et accorder une attention particulière aux plus vulnérables (art. 74, paragr. e); ils devraient également être évolutifs et adaptatifs en tenant compte progressivement des spécificités des pays et des situations, des nouvelles méthodes et des cadres et processus existants (art. 74, paragr. f); ils devraient être rigoureux, basés sur des preuves scientifiques, soutenus par des évaluations menées par les pays et des données de qualité, accessibles, actualisées, fiables et ventilées selon le revenu, le sexe, l'âge, la race, l'appartenance ethnique, le statut migratoire, le handicap, la situation géographique (art. 74, paragr. g); ils devraient bénéficier d'un appui accru en matière de renforcement des capacités pour les pays en développement en matière de systèmes de collecte de données et d'évaluation au niveau national (art. 74, paragr. h), ainsi que du soutien actif de l'ONU et d'autres institutions multilatérales (art. 74, paragr. i).

L'examen de l'agenda 2030 pour le développement durable se réalise aux niveaux national, régional et mondial.

Au niveau national, on s'attend à ce que chaque État Membre élabore, sur une base volontaire, un rapport écrit, mieux connu sous l'appellation rapport national volontaire ou examen national volontaire (ENV), retraçant les progrès réalisés aux niveaux national et local par rapport à un cadre de référence national adapté et ambitieux, et en s'appuyant sur les mécanismes nationaux existants en matière de planification et d'examen (art. 78-79). Dans cet exercice, il est attendu des autorités publiques d'assurer un processus inclusif et participatif en impliquant toutes les parties prenantes concernées.

Au niveau régional, en se basant sur les processus d'examen au niveau national, les examens régionaux ont un rôle crucial à jouer au niveau de la coordination des ENV et la contribution au processus de suivi et d'examen au niveau mondial. Ceci doit se réaliser à travers notamment : (i) la transmission de connaissances entre pairs, notamment à travers les ENV; (ii) l'échange de meilleures pratiques et l'enseignement mutuel; (iii) la discussion sur les cibles communes; (iv) l'identification des goulots d'étranglement et la mobilisation des énergies; (v) la transmission des résultats des rapports régionaux au forum politique de haut niveau sous forme agréée (Nations Unies, 2015 : article 80).

Le niveau mondial est assuré par le Forum politique de haut niveau sur le développement durable (FPHN) qui supervise l'ensemble du processus de suivi et d'examen en collaboration avec l'Assemblée générale, le Conseil Économique et Social (CES) et d'autres instances concernées.

Constituant la plateforme centrale d'examen et de suivi au niveau global, le FPHN se réunit tous les quatre ans sous la direction de l'Assemblée générale, au niveau des chefs d'État et de gouvernement, et annuellement, sous la direction du CES. Les objectifs/mandats du FPHN sont : (i) apprécier et promouvoir la cohérence de l'ensemble du système; (ii) renforcer le partenariat entre les différentes parties prenantes; (iii) faire un bilan régulier des réalisations afin d'évaluer les progrès accomplis; (iv) prendre connaissance des contraintes, des défis et des opportunités; (v) faciliter la mise en commun des expériences; (vi) tirer les enseignements de la mise en œuvre pour donner les orientations et recommandations nécessaires pour garantir une meilleure réalisation des résultats (Nations Unies, 2013; Nations Unies, 2015 : articles 82-84). Le FPHN se base ainsi sur les ENV et des rapports écrits des États, des rapports régionaux, le rapport annuel du Secrétaire général sur les progrès des ODD, l'examen annuel de l'objectif 17 ainsi que sur des examens thématiques.

## L'évaluation dans l'agenda 2030 pour le développement durable et les examens nationaux volontaires : des conditions indispensables pour leur réalisation

Bien que le programme de développement durable à l'horizon 2030 reconnaisse explicitement l'importance de l'évaluation (Nations Unies, 2015 : paragr. 72 et 74(g)), le travail à l'heure actuelle se focalise beaucoup plus sur les problématiques de collecte des données et le choix des indicateurs. Ce travail laborieux constitue **une condition nécessaire pour assurer une meilleure compréhension des réalisations accomplies au niveau de chaque ODD, mais reste insuffisant** pour apporter des réponses aux questions classiques de l'évaluation et savoir « si nous faisons les choses correctement » ou « si on a fait les bonnes choses » et « si on aurait dû faire autrement et mieux ».

Dans la pratique, les examens nationaux volontaires élaborés pendant les six premières années de l'entrée en vigueur de l'agenda 2030 pour le développement durable s'apparentent beaucoup plus à des exercices d'exhibition décrivant les réalisations dans les différents domaines des ODD et des

étapes franchies dans la sensibilisation, l'intégration des ODD au niveau national et la mise en place des cadres institutionnels pour le pilotage des différents processus **qu'à une réelle revue des réalisations selon les cibles et les indicateurs établis.**

Initialement, il a été prévu que durant les premières années, le processus de suivi et d'examen se concentre sur le progrès réalisé dans l'intégration des ODD dans les plans, stratégies et politiques au niveau national, ainsi que sur leur adaptation aux conjonctures nationales et l'ajustement de dispositifs institutionnels pertinents. Durant les années suivantes, la revue porterait sur la réalisation concrète des objectifs. Cela impliquerait le suivi du progrès par rapport aux cibles et aux indicateurs, l'évaluation des politiques et programmes. Il est ainsi attendu que le FPHN soit plus exigeant en matière de suivi et d'évaluation des contributions des pays dans la réalisation des ODD selon les cibles et les indicateurs établis et contextualisés. La rigueur exigée de cet exercice semble largement tributaire d'un certain nombre de préalables. La liste suivante présente ceux que nous estimons essentiels.

### **La domestication des ODD et leur adaptation au contexte national à travers la définition de cibles spécifiques, pertinentes et réalisables, tout en étant ambitieux, en fonction des priorités et des conditions initiales de développement du pays**

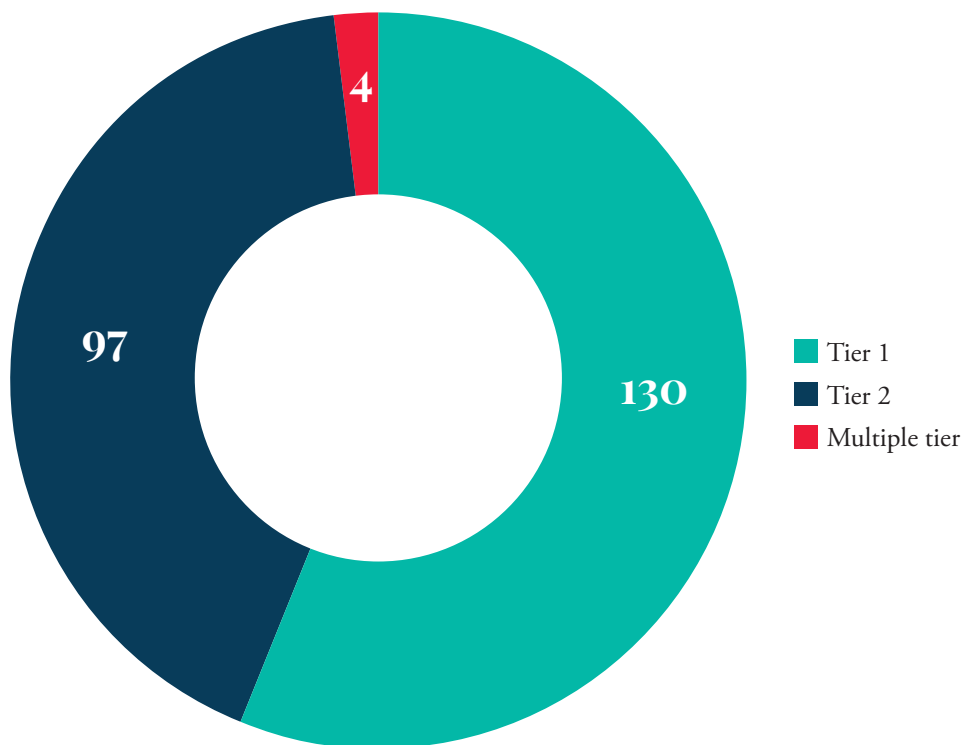
Tirant enseignement des OMD et afin d'éviter le manque d'appropriation nationale des ODD globaux, il semble indispensable de procéder à l'adaptation de ces objectifs aux particularités nationales. Cependant, une réelle appropriation nationale reste largement tributaire de la volonté et de la capacité du pays à élaborer une stratégie nationale à l'horizon 2030 sur le développement durable, qui intègre les différentes dimensions (économique, sociale, environnementale et de gouvernance), les différentes régions, localités et parties prenantes, tout en tenant compte des spécificités et des conditions initiales de développement du pays. L'élaboration d'une telle stratégie semble constituer un préalable à une évaluation rigoureuse susceptible d'expliquer le pourquoi et le comment des changements observés des ODD dans un pays et leur importance pour les différentes parties prenantes et notamment les personnes laissées pour compte.

### **L'adoption au niveau national des normes statistiques internationales**

L'écart entre les normes statistiques internationales et nationales pose un vrai défi, d'autant plus que le suivi mondial doit être basé essentiellement sur des données nationales comparables et standardisées. Cependant, les données collectées au niveau national ne sont pas toujours comparables ou même manquantes pour certains indicateurs. Un grand nombre de pays en développement ne disposeront pas de données de référence pour plusieurs indicateurs mondiaux des ODD, surtout dans les nouveaux domaines.

## La disponibilité des données, notamment de référence, et l'établissement des méthodologies pour l'essentiel des indicateurs

Selon la dernière mise à jour du groupe des experts des Nations Unies et de l'extérieur (GENUE) réalisée le 29 mars 2021<sup>1</sup>, sur les 231<sup>2</sup> indicateurs uniques établis, seulement 130 indicateurs appartiennent au premier groupe pour lequel il existe des méthodes statistiques éprouvées et des données mondiales disponibles sur une base régulière. Les autres indicateurs sont classés ou bien dans le deuxième groupe qui est composé de 97 indicateurs associés à des méthodes statistiques définies, mais avec peu de données disponibles, ou non encore attribués à un groupe (4 indicateurs). Le travail du GENUE se poursuivra en explorant de nouvelles sources de données et technologies pour la collecte des données en partenariat avec la société civile, les milieux universitaires et le secteur privé, et en établissant des références concernant les indicateurs pour lesquels les données ne sont toujours pas disponibles.



**Figure 2**

Classification des indicateurs ODD.

Source : À partir des données du Département des affaires économiques et sociales (Consulté le 29 mars 2021).

1. [https://unstats.un.org/sdgs/files/Tier%20Classification%20of%20SDG%20Indicators\\_29%20Mar%202021\\_web.pdf](https://unstats.un.org/sdgs/files/Tier%20Classification%20of%20SDG%20Indicators_29%20Mar%202021_web.pdf)

2. Le nombre d'indicateurs mondiaux uniques relatifs aux ODD adoptés par la Commission de statistique de l'ONU, et actualisés et approuvés lors de la 52<sup>e</sup> session de la Commission de statistique en mars 2021.

## La capacité d'aller au-delà du suivi des indicateurs et de gérer la complexité du système des ODD

L'explication du changement produit et son impact réel sur les différentes parties prenantes reste largement tributaire de la capacité du cadre de suivi et d'examen à aller au-delà du suivi des indicateurs établis pour apporter des jugements éclairés sur la pertinence, l'efficacité, l'impact, la cohérence et la durabilité<sup>3</sup> des ODD d'une part, et des politiques et programmes afférents d'autre part. Cependant, ces jugements ne seront pas faciles à mener vu la complexité du système des ODD. En effet, le nombre élevé de ces derniers, des cibles afférentes et leurs interliaisons d'une part, et la multitude de parties prenantes à intérêts souvent divergents et leur interaction ainsi que l'horizon temporel relativement long d'autre part, constituent autant d'éléments qui renforcent l'incertitude et qui confirment la complexité du système des ODD. Cette complexité signifie que le lien de causalité linéaire et direct entre une politique ou un programme donné et la réalisation d'un ODD ne peut pas être parfaitement et facilement confirmé. Cette situation est de nature à rendre la tâche de l'évaluateur plus difficile pour **dégager l'effet propre d'une intervention sur le résultat obtenu**. Elle nécessitera ainsi l'utilisation de méthodologies issues de la science de la complexité et la mise en place d'une gestion adaptative capable via la connaissance produite par le système de suivi et d'évaluation, afin d'assurer une meilleure compréhension de la réalité et d'explorer de nouvelles voies et de solutions alternatives *out of the box* pour la réalisation des objectifs fixés et des cibles prévues en utilisant des moyens progressivement innovants. Des solutions capables de ne pas reproduire dans le futur les problèmes que nous avons aujourd'hui, comme le suggère Albert Einstein dans sa citation « on ne résout pas un problème avec les modes de pensée qui l'ont engendré. »

## Conclusion

Une revue rigoureuse des réalisations des ODD selon les cibles et indicateurs établis reste largement tributaire de : i) la domestication des ODD et leur adaptation au contexte national; ii) l'adoption au niveau national des normes statistiques internationales; iii) la disponibilité des données, notamment de référence, et l'établissement des méthodologies pour l'essentiel des indicateurs; iv) la capacité d'aller au-delà du suivi des indicateurs et de gérer la complexité du système des ODD. Il est ainsi attendu que le FPHN soit plus exigeant en matière de suivi et d'évaluation des contributions des pays dans la réalisation des ODD.

---

3. Les six critères d'évaluation du Comité d'aide au développement (CAD) de l'Organisation de coopération et de développement économique (OCDE) selon la dernière révision du réseau CAD sur l'évaluation du développement afin d'assurer une meilleure adaptation des critères au nouveau paysage du développement et à l'Agenda 2030 (décembre 2019). <https://www.oecd.org/fr/cad/evaluation/criteres-adaptees-evaluation-dec-2019.pdf>

## Pour aller plus loin

Forum politique de haut niveau sur le développement durable. (2021). *Manuel de préparation des examens nationaux volontaires*. [https://sustainabledevelopment.un.org/content/documents/27054Handbook\\_2021\\_FR.pdf](https://sustainabledevelopment.un.org/content/documents/27054Handbook_2021_FR.pdf)

Ce manuel est mis à jour annuellement et il est destiné à la préparation et à la présentation des Examens nationaux volontaires (ENV). Pour consulter ce manuel en ligne ou les dernières informations sur les Examens nationaux volontaires, rendez-vous respectivement sur [https://sustainabledevelopment.un.org/content/documents/27054Handbook\\_2021\\_FR.pdf](https://sustainabledevelopment.un.org/content/documents/27054Handbook_2021_FR.pdf) et <https://sustainabledevelopment.un.org/vnrs/>

Pour consulter les dernières mises à jour du cadre des indicateurs des ODD et des travaux de la Commission de statistique des Nations Unies, rendez-vous sur <https://unstats.un.org/sdgs/iaeg-sdgs/tier-classification/>

Pour avoir une idée sur les ODD atteignables à l'horizon 2030 et ceux non atteignables au même horizon, ainsi que sur les effets de la pandémie de COVID-19 sur les ODD, veuillez consulter le rapport 2021 sur les ODD. [https://unstats.un.org/sdgs/report/2021/The-Sustainable-Development-Goals-Report-2021\\_French.pdf](https://unstats.un.org/sdgs/report/2021/The-Sustainable-Development-Goals-Report-2021_French.pdf)

## Références bibliographiques

Bamberger, M., Segone, M. et Tateossian, F. (2017). *Évaluations axées sur l'équité et sensibles au genre : Assurer le suivi des objectifs de développement durable selon la perspective « aucun-laissé-pour-compte »*. ONU Femmes, EvalGender+ et IOCE. <https://www2.unwomen.org/-/media/field%20office%20morocco/documents/publications/2017/10/fr-evaluating-sdg-web.pdf?la=fr&vs=2152>

Département des Affaires Économiques et Sociales des Nations Unies. (2016). *Annexe : Propositions de directives communes d'application volontaire sur les rapports aux fins des examens nationaux volontaires au sein du forum politique de haut niveau*. [https://sustainabledevelopment.un.org/content/documents/11843Voluntary\\_guidelines\\_VNRs\\_French.pdf](https://sustainabledevelopment.un.org/content/documents/11843Voluntary_guidelines_VNRs_French.pdf)

Département des Affaires Économiques et Sociales des Nations Unies. (2021). *Tier Classification for Global SDG Indicators as of 29 March 2021*. [https://unstats.un.org/sdgs/files/Tier%20Classification%20of%20SDG%20Indicators\\_29%20Mar%202021\\_web.pdf](https://unstats.un.org/sdgs/files/Tier%20Classification%20of%20SDG%20Indicators_29%20Mar%202021_web.pdf)

Forum politique de haut niveau sur le développement durable. (2021). *Manuel de préparation des examens nationaux volontaires*. [https://sustainabledevelopment.un.org/content/documents/27054Handbook\\_2021\\_FR.pdf](https://sustainabledevelopment.un.org/content/documents/27054Handbook_2021_FR.pdf)

Ghouil, M. (2016). *L'agenda 2030 pour le développement durable. Genèse, principes et processus d'élaboration des Objectifs du Développement Durable (ODD)*. Bureau des Nations Unies. [http://www.onu-tn.org/Publications/Articles/28\\_NEW\\_Agenda\\_2030\\_pour\\_le\\_developpement\\_durable\\_genese\\_principes\\_et\\_processus](http://www.onu-tn.org/Publications/Articles/28_NEW_Agenda_2030_pour_le_developpement_durable_genese_principes_et_processus)

Ghouil, M. (2017). *Examen, suivi & évaluation des objectifs de développement durable : nouveautés, spécificités et défis*. Bureau des Nations Unies. <http://www.onu-tn.org/uploads/articles/14930307090.pdf>



- Nations Unies. (2017a). *Jalons essentiels sur la voie d'un suivi et d'un examen cohérents, efficient et inclusif au niveau mondial : Rapport du Secrétaire général (publication n° 16-00241)*. <https://ncdalliance.org/fr/resources/jalons-essentiels-sur-la-voie-d'un-suivi-et-d'un-examen-coherents-efficients-et-inclusifs-au-niveau-mondial>
- Nations Unies. (2020). *Rapport du Groupe d'experts des Nations Unies et de l'extérieur chargé des indicateurs relatifs aux objectifs de développement durable (publication n° 19-22255)*. <https://unstats.un.org/unsd/statcom/51st-session/documents/2020-2-SDG-IAEG-F.pdf>
- Nations Unies. (2021). *Rapport sur les Objectifs de Développement Durable 2021*. [https://unstats.un.org/sdgs/report/2021/The-Sustainable-Development-Goals-Report-2021\\_French.pdf](https://unstats.un.org/sdgs/report/2021/The-Sustainable-Development-Goals-Report-2021_French.pdf)
- Résolution 67/290 de l'Assemblée générale, *Format and organizational aspects of the high-level political forum on sustainable development*, A/RES/67/290 (23 août 2013) [https://www.un.org/ga/search/view\\_doc.asp?symbol=A/RES/67/290&Lang=E](https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/67/290&Lang=E)
- Résolution 70/1 de l'Assemblée générale, *Transformer notre monde : le Programme de développement durable à l'horizon 2030*, A/RES/70/1 (21 octobre 2015) [https://unctad.org/system/files/official-document/ares70d1\\_fr.pdf](https://unctad.org/system/files/official-document/ares70d1_fr.pdf)
- Résolution 70/299 de l'Assemblée générale, *Suivi et examen de la mise en œuvre du Programme de développement durable à l'horizon 2030 au niveau mondial*, A/RES/70/299 (18 août 2016) <https://undocs.org/fr/A/RES/70/299>
- Résolution 71/313 de l'Assemblée générale, *Travaux de la Commission de statistique sur le Programme de développement durable à l'horizon 2030*, A/RES/71/313 (10 juillet 2017) [http://ggim.un.org/documents/A\\_Res\\_71\\_313\\_f.pdf](http://ggim.un.org/documents/A_Res_71_313_f.pdf)
- The United Nations Institute for Training and Research (UNITAR-GNUD). (2015). *Se préparer à l'action : le Programme à l'horizon 2030 et les ODD* [dossier des séminaires nationaux, module 5, diapositives 2].



# Enjeux et principes de l'évaluation des mouvements et politiques migratoires

*Diana Cartier*

*Christophe Franzetti*

## Introduction

La migration est un phénomène global qui a un impact à tous les niveaux de la société, de l'économie, de la politique, de la culture et du marché du travail. Ces dernières années, plusieurs événements ont mis la réalité migratoire sous les projecteurs, en particulier ceux liés à des crises humanitaires.

Du point de vue politique, un changement historique s'est produit à l'échelle mondiale lorsque les États Membres des Nations Unies ont conclu en 2018 deux pactes mondiaux sur les migrations et les déplacements à l'échelle internationale : le Pacte mondial pour des migrations sûres, ordonnées et régulières et le Pacte mondial sur les réfugiés<sup>1</sup>. Leur adoption marque l'aboutissement d'efforts déployés depuis des décennies par les États, les organisations internationales, les organisations de la société civile et d'autres acteurs pour améliorer les modalités de la gestion des migrations à l'échelle internationale (OIM, 2020).

Sur les plans social et économique, les crises humanitaires de ces dernières années ont provoqué des déplacements de millions de personnes, en raison d'un conflit ou d'une grave instabilité économique et politique, provoquant de grandes souffrances et de graves traumatismes ainsi que la perte de nombreuses vies humaines. On a également pu constater l'augmentation des dommages liés aux effets des changements climatiques et environnementaux, et leurs impacts sur la mobilité humaine.

En 2019, par rapport à l'année précédente, 33,4 millions de personnes nouvellement déplacées de plus à l'intérieur de leurs pays ont été recensées. De ce nombre, 8,5 millions de personnes ont été déplacées en raison de conflits ou de violences sociales et/ou politiques, et 24,9 millions à cause de désastres naturels (IDMC, 2020). Si l'on considère les personnes encore déplacées depuis plusieurs années, le total de personnes déplacées atteint 50,8 millions (IDMC, 2020). Il est intéressant de relever qu'au cours des dix dernières années, le nombre annuel de nouveaux déplacements dus à des

---

1. Le 19 septembre 2016, les chefs d'État et de gouvernement des 193 États Membres de l'ONU se sont réunis à l'Assemblée générale des Nations Unies pour discuter de questions liées à la migration et aux réfugiés à l'échelle mondiale. L'adoption de la Déclaration de New York pour les réfugiés et les migrants a souligné la nécessité d'envisager la migration de manière globale. Par conséquent, les États Membres de l'ONU ont convenu de coopérer à l'élaboration d'un Pacte mondial pour des migrations sûres, ordonnées et régulières, qui a été adopté lors d'une conférence intergouvernementale sur les migrations internationales les 10 et 11 décembre 2018 au Maroc. La Déclaration de New York pour les réfugiés et les migrants a également initié un processus de négociation distinct pour le Pacte mondial sur les réfugiés.

catastrophes naturelles est supérieur à ceux dus à un conflit ou à la violence, avec davantage de pays touchés : en 2018, 144 pays étaient affectés du fait de catastrophes naturelles, contre 55 en raison d'un conflit ou de violence (IDMC, 2020).

D'une manière plus globale, une augmentation des migrations internationales a été constatée avec une estimation de 272 millions de personnes migrantes à travers le monde, représentant 3,5 % de la population mondiale, dont près des deux tiers sont des travailleurs migrants. L'ampleur et le rythme des migrations internationales sont toutefois difficiles à prédire, et leur gestion compliquée à planifier, du fait de la nature imprévisible des conflits et des désastres naturels, et à l'aspect clandestin de la migration irrégulière (OIM, 2020).

À partir de la littérature et de nos expériences, ce chapitre présente cinq principes qui devraient guider les évaluations portant sur les politiques de migration internationale et sur les déplacements de population pour faire face aux enjeux spécifiques de la migration.

## **Enjeux et principes de l'évaluation de mouvements et politiques migratoires**

Au-delà du contexte migratoire décrit dans l'introduction, plusieurs spécificités doivent être prises en compte lors de l'évaluation de mouvements et politiques migratoires, telles que les enjeux politiques et socio-économiques, ainsi que l'aspect souverain des politiques migratoires comme relevé dans le Pacte mondial sur les migrations.

Toutefois, il existe différentes définitions de la migration et des personnes migrantes qui tiennent compte de facteurs géographiques, juridiques, politiques ou temporels. Ces différences de définitions et de méthodes de collecte de données d'un pays à l'autre rendent parfois difficiles les comparaisons basées sur les statistiques nationales et d'éventuelles généralisations au niveau mondial.

### **L'importance des sources d'information face à une perception négative du phénomène migratoire**

La perception du phénomène migratoire est souvent controversée et la migration fortement politisée avec ses relents populistes, ses slogans xénophobes et son utilisation par le public, par les médias et par les groupes politiques pour juger de la performance d'un gouvernement (Chappell et Laczko, 2011). Dans de telles circonstances, il est important de bien s'informer et de vérifier les différentes sources d'informations pour s'assurer d'une évaluation objective, non partisane et fondée sur un ensemble de données factuelles.

Beaucoup d'efforts ont été menés ces dernières années en ce qui concerne les données sur la migration afin de documenter les progrès réalisés dans le cadre de l'Agenda 2030 et plus récemment avec l'élaboration du pacte mondial. Les bureaux nationaux de statistiques peuvent être à cet égard une source de données riches et fiables ainsi que les bases de données élaborées par des organisations internationales qui traitent de la migration (par exemple BIT, OIM, DAES, HCR, au sein du système des Nations Unies<sup>2</sup>). Les déplacements internes et les flux migratoires dans des contextes de crises sont également bien documentés en termes statistiques.

---

2. BIT : Bureau International du Travail; OIM : Organisation internationale pour les migrations; DAES : Département des affaires économiques et sociales; HCR : Le Haut Commissariat des Nations Unies pour les réfugiés.

## **Le Pacte mondial sur les migrations comme outil d'analyse pour l'élaboration et l'évaluation des politiques des gouvernements face aux migrations**

Il est important de bien cerner les politiques des gouvernements sur le sujet de la migration. Il est fort possible qu'il n'existe pas de politique globale sur la gestion de la migration, étant incorporée dans diverses législations ou stratégies gouvernementales relevant de différents ministères. Le Pacte mondial sur les migrations est intéressant à ce niveau pour son approche globale et cohérente du phénomène migratoire qui peut servir d'outil d'analyse et qui peut permettre une étude comparative selon diverses composantes.

Le Programme de développement durable à l'horizon 2030 adopté en 2015, contrairement aux objectifs du Millénaire pour le développement de la Déclaration du Millénaire adopté en 2000, fait référence aux personnes migrantes, réfugiées et déplacées internes, et reconnaît que la migration est aussi une composante du développement (OIM, 2018). Il souligne, entre autres, que les femmes, les hommes et les enfants migrants forment un groupe vulnérable qu'il faut aussi protéger en tant qu'agents du développement.

La cible 10.7 de l'ODD visant à « réduire les inégalités dans les pays et d'un pays à l'autre » appelle à « faciliter la migration et la mobilité de façon ordonnée, sans danger, régulière et responsable, notamment par la mise en œuvre de politiques de migration planifiées et bien gérées ». En outre, d'autres cibles se réfèrent directement à la migration ou mentionnent la migration comme une question transversale qu'il faut prendre en compte (OIM, 2018).

Les indicateurs de gouvernance des migrations (IGM) de l'Organisation internationale pour les migrations peuvent être utilisés comme un outil de diagnostic dans le choix des cibles. Élaborés avec *l'Economist Intelligence Unit* (EIU), les IGM constituent un outil permettant aux pays d'évaluer leur gouvernance des migrations relative à la cible 10.7 et d'approfondir le dialogue sur celle-ci.

Le Pacte mondial sur les migrations a apporté une nouvelle impulsion pour mettre en œuvre les aspects des objectifs du développement durable (ODD) liés à la migration, et a apporté davantage de précisions puisqu'il est compatible avec la cible 10.7 selon laquelle les États Membres s'engagent à coopérer à l'échelle internationale en vue de faciliter la migration de façon ordonnée, sûre, régulière et responsable.

Le Pacte mondial sur les migrations n'est pas contraignant, mais se présente comme une feuille de route détaillée permettant l'amélioration de la gouvernance migratoire et la définition de son cadre d'évaluation.

Au niveau national, il est important d'analyser dans quelle mesure la migration est bien incorporée dans les politiques du pays. Il est probable que s'il n'y a pas de politique proprement liée à la migration, d'autres politiques et cadres législatifs influencent les approches et les besoins des personnes migrantes – par exemple sur le marché du travail ou l'éducation, et leur lien avec l'immigration dans le pays (Czaika et de Haas, 2015).

## **L'individu au centre de la prise en compte des motifs migratoires et des facteurs de vulnérabilité spécifiques à certaines populations**

Étant donné les diverses raisons qui influencent une personne à migrer, la prise en compte des objectifs qui la motivent lors du processus d'évaluation facilitera la compréhension de sa situation et de ses besoins. Elle permettra d'analyser le type de soutien approprié, en considérant également des contraintes

liées au lieu où se trouve la personne migrante. La pluralité des dimensions du phénomène migratoire doit ainsi être intégrée dans l'évaluation, y compris les thèmes transversaux comme le genre, les droits de l'homme, la protection des personnes vulnérables ou l'environnement.

Au-delà de son intention de promouvoir des politiques migratoires ordonnées et responsables, la communauté internationale est également préoccupée par les situations précaires et les dangers vécus par les personnes migrantes victimes de crises et de conflits, y compris dans leurs tentatives d'atteindre le pays de destination de leur choix. Si les crises touchent indifféremment les nationaux et les non-nationaux, les personnes migrantes ont parfois plus de mal à faire face à leurs effets dévastateurs en raison des vulnérabilités liées à leur statut juridique, économique et social et aux obstacles liés à la langue ou à leur culture. Par conséquent, dans le cadre d'une évaluation, les effets sur les personnes migrantes pourraient être étudiés, y compris l'impact environnemental ou social d'une intervention sur ces personnes.

Comme outil d'évaluation pour les organismes travaillant dans le cadre de crises migratoires, le Cadre opérationnel en cas de crise migratoire (COCM) peut être utile. Ce cadre d'analyse et de planification est utilisé en soutien aux gouvernements afin qu'ils puissent mieux se préparer aux crises migratoires, et aider à s'en relever (OIM, 2018). Le Cadre opérationnel décrit tous les secteurs d'interventions qui pourraient faire partie de la réponse à la crise en s'appuyant sur les activités humanitaires traditionnelles et des activités de consolidation de la paix, de transition et de développement ainsi que la gestion des migrations. Ceci permet aussi d'évaluer l'impact de la crise migratoire au niveau individuel et sociétal.

Le Pacte mondial sur les migrations, adopté en 2018, considère également que la situation des personnes migrantes prises dans une crise est un élément essentiel de la réalisation des objectifs 2 et 7, qui visent respectivement à limiter au minimum les facteurs de migration forcée et à réduire la vulnérabilité de ces personnes.

Concernant les thèmes transversaux, par exemple pour intégrer le genre dans toute évaluation, certaines spécificités liées à la migration doivent être prises en compte. Les femmes et les filles migrantes peuvent être confrontées à des obstacles sexospécifiques et discriminatoires, entravant leur mobilité et mettant en danger leur personne, surtout lors de migrations irrégulières, en s'exposant à un risque accru d'exploitation et de violence sexuelle. Ces vulnérabilités différenciées ne sont pas forcément détaillées dans la gouvernance des migrations, et intégrer l'analyse des besoins respectifs des femmes, des hommes, des filles et des garçons migrants dans les évaluations serait un pas important dans la reconnaissance du phénomène du genre, et autres aspects transversaux comme les handicaps, dans les politiques et les programmes futurs sur la migration (OIM, 2018).

### **La prise en compte de la voix des personnes migrantes à travers des évaluations participatives**

Une évaluation participative qui permet de faire entendre la voix des personnes migrantes et à l'inclure dans le processus aussi tôt que possible permettra une évaluation compréhensive et ciblée sur les besoins, les attentes et les espoirs des migrants, au-delà de l'application des politiques migratoires par les gouvernements.

S'il est primordial d'incorporer la voix des migrants dans un processus d'évaluation, comprendre les perceptions et les attentes des communautés hôtes confrontées aux phénomènes migratoires est tout aussi important. Un tel dialogue avec les parties concernées lors d'une évaluation permettra de proposer des analyses pertinentes sur les réalités xénophobes, sur l'intégration des personnes migrantes ou sur les possibilités d'approches positives de la migration en tant que contribution à l'économie et à la richesse des échanges culturels.

## **L'importance de la planification d'une évaluation dès la conception de l'intervention et tout au long de son cycle de vie**

La préparation d'une évaluation touchant à la migration commence lors de la conception du programme ou de l'intervention afin de cerner proprement le contexte dans lequel se développe le phénomène migratoire, son impact, et afin de lui attribuer son juste rôle.

Le principal défi pour mesurer l'effet des politiques ou des interventions liées à la migration est d'identifier leur contribution par rapport à d'autres facteurs d'influence possibles, que ce soit en matière de développement économique, de gestion ordonnée des migrations, d'intégration sociale ou dans le respect des droits de l'homme.

Par exemple, un pays peut décider de mettre en œuvre une politique visant à encourager ses nationaux à migrer vers un pays tiers dans le but d'augmenter le revenu national et l'apport de devises étrangères grâce aux envois de fonds. Toutefois, des effets collatéraux négatifs peuvent se produire si ce même gouvernement n'est pas à même d'aider ces personnes migrantes dans le pays de destination et de les protéger d'abus dans leur travail ou du fait de leur situation précaire. Une réflexion sur la conduite d'une évaluation de cette politique et sa mise en œuvre dès sa conception permettra d'intégrer l'ensemble des effets possibles, leur origine et comment les mesurer (Global Migration Group, 2010).

## **Conclusion**

La migration continuera d'être un phénomène important dans les politiques gouvernementales étant donné son impact à tous les niveaux de la société, de l'économie, de la politique, de la culture et du marché du travail. L'évaluateur déchiffre la complexité, les enjeux et l'impact de la migration pour proposer une analyse plus approfondie, incluant les voix des migrants et en utilisant les outils mis à leur disposition tels que le Pacte mondial ou le Cadre opérationnel.

## **Pour aller plus loin**

De l'information sur le Pacte mondial sur les migrations peut être trouvée sur ce site : <https://migrationdataportal.org/fr/global-compact-for-migration>

Sur ce site, les lecteurs peuvent trouver des documents thématiques contribuant au Pacte mondial : <https://www.iom.int/iom-thematic-papers>

Sur ce site, un guide pratique est accessible sur la migration et le Programme 2030 : <https://publications.iom.int/books/la-migration-et-le-programme-2030-un-guide-lusage-des-praticiens>

Sur ce site, des profils de migration par pays peuvent être trouvés et utilisés par les évaluateurs : <http://www.gfmd.org/pfp/policy-tools/migration-profiles/repository>

D'autres données sont disponibles (en anglais) sur la migration sur ces sites :

<https://globalmigrationgroup.org/content/data-and-research-data-sources>

<https://www.iom.int/fr/termes-cles-de-la-migration>

<https://www.internal-displacement.org/>

ODD : <https://www.un.org/sustainabledevelopment/fr/inequality/>

## Références bibliographiques

- Assemblée générale des Nations Unies, 73<sup>e</sup> session, *Rapport du Haut-Commissaire des Nations Unies pour les réfugiés. Deuxième partie. Pacte mondial sur les réfugiés*, Supplément n°(A/73/12 (Part II) (2018), [https://undocs.org/fr/A/73/12\(PartII\)](https://undocs.org/fr/A/73/12(PartII))
- Chappell, L. et Laczko, F. (2011, 8 juin). What are we Really Achieving? Building an Evaluation Culture in Migration and Development. *The Online Journal of the Migration Policy Institute*. <https://www.migrationpolicy.org/article/what-are-we-really-achieving-building-evaluation-culture-migration-and-development>
- Czaika, M. et de Haas, H. (2015). Evaluating Migration Policy Effectiveness. Dans A. Triandafyllidou et I. Isaakyan (dir.), *Routledge Handbook of Immigrant and Refugee Studies*. Routledge. <https://heindehaas.files.wordpress.com/2015/05/2015-czaika-and-de-haas-evaluating-migration-policy-effectiveness.pdf>
- Global Migration Group. (2010). *Mainstreaming Migration into Development Planning : A handbook for policy-makers and practitioners*. OIM.
- Internal Displacement Monitoring Centre (IDMC). (2020). *Global report on Internal Displacement 2020 Summary*. <https://www.internal-displacement.org/global-report/grid2020/>
- OIM. (2018). *La migration et le Programme 2030 : Un guide à l'usage des praticiens*. <https://publications.iom.int/books/la-migration-et-le-programme-2030-un-guide-lusage-des-praticiens>
- OIM. (2020). *État de la migration dans le Monde 2020*. <https://www.iom.int/wmr/2020?language=1450#content>
- Résolution 55/2 de l'Assemblée générale, *Déclaration du Millénaire*, A/55/2 (13 septembre 2000) <https://undocs.org/en/A/RES/55/2>
- Résolution 70/1 de l'Assemblée générale, *Transformer notre monde : le Programme de développement durable à l'horizon 2030*, A/RES/70/1 (21 octobre 2015) <https://undocs.org/fr/A/RES/70/1>
- Résolution 71/1 de l'Assemblée générale, *Déclaration de New York pour les réfugiés et les migrants*, A/RES/71/1 (3 octobre 2016), <https://undocs.org/fr/A/RES/71/1>
- Résolution 73/195 de l'Assemblée générale, *Pacte mondial pour des migrations sûres, ordonnées et régulières*, A/RES/73/195 (11 janvier 2019), <https://undocs.org/fr/A/RES/73/195>



# Évaluation des projets de coopération internationale sur l'adaptation aux changements climatiques : enjeux et défis

*Nicolas Montibert*

## Introduction

Les enjeux de développement liés aux changements climatiques ne cessent de croître dans la plupart des pays du monde, en particulier dans les pays à faibles revenus, qui n'ont pas les moyens financiers et techniques de s'adapter. Les phénomènes climatiques extrêmes, de plus en plus fréquents et intenses, exacerbent les défis déjà importants relatifs à la sécurité alimentaire, à l'accès à l'eau potable, à la santé des populations vulnérables, à la gestion des risques et des désastres ou à l'égalité entre les hommes et les femmes; ces dernières étant davantage affectées par les effets des changements climatiques. Pour faire face à cette réalité, de plus en plus de projets de développement conçoivent des stratégies, spécifiques ou transversales, visant à renforcer l'efficacité et la durabilité de leurs interventions face aux variations du climat. Les méthodologies utilisées pour l'évaluation de ces interventions sont généralement comparables à celles des autres thématiques, certaines intégrant également une analyse globale de l'influence des conditions climatiques sur le niveau d'atteinte des résultats.

Pourtant, l'évaluation rigoureuse d'interventions visant à renforcer l'adaptation et la résilience des populations aux changements climatiques exige une approche plus complexe qui révèle plusieurs défis méthodologiques et techniques. Elle exige, en effet, une bonne compréhension du contexte biophysique dans son ensemble, et doit aussi tenir compte des facteurs de vulnérabilité des populations cibles, notamment sur le plan socio-économique, qu'il s'agit, de surcroît, d'appréhender selon une perspective holistique et dynamique. L'évaluation de ce type d'interventions doit, par conséquent, considérer non seulement l'évolution des conditions climatiques, mais aussi celle de ses impacts, qui ne dépendent pas uniquement de l'intensité des phénomènes climatiques, mais également du degré d'exposition et de vulnérabilité de la population considérée. Développer une méthodologie d'évaluation qui prenne en compte l'influence de ces multiples facteurs suppose donc de contourner plusieurs défis récurrents dans les pays en développement. Ce chapitre propose d'aborder un certain nombre de ces défis concrets à partir de l'expérience de l'auteur.

## **Défis pour l'évaluation des projets d'adaptation et de résilience aux changements climatiques**

Les défis présentés ci-dessous sont présentés par ordre d'importance, d'après l'expérience de l'auteur, même si cette hiérarchie est variable et dépend étroitement des pays d'intervention, de la qualité du plan de suivi-évaluation et du budget qui lui est affecté.

### **Difficulté d'accès et manque de précision des données climatiques**

La difficulté d'accès et le manque de précision des données climatiques nationales, qui constituent une information de base pour la compréhension du contexte d'intervention et des facteurs externes ayant pu freiner ou faciliter l'atteinte des résultats escomptés, restent des défis importants pour la qualité de l'analyse et de l'évaluation. Notre expérience révèle que la disponibilité même des données météorologiques reste une difficulté importante dans la plupart des pays à faibles à revenus, dont les réseaux de stations météorologiques restent peu denses et pas toujours pleinement fonctionnels, en particulier dans les zones rurales isolées d'Afrique ou d'Haïti (PNUD, 2016). La qualité et la fiabilité des informations disponibles posent aussi parfois problème, d'autant que leur précision se limite souvent à l'échelle régionale, alors que les changements climatiques ont tendance à accentuer les événements météorologiques extrêmes et localisés (IPCC, 2021), tels que les poches de sécheresse, les précipitations violentes, les épisodes de grêle ou de gel, les vents violents, etc. Ainsi, des moyennes mensuelles ou annuelles relativement conformes aux normales saisonnières peuvent dissimuler de fortes variations dans la répartition des précipitations, ainsi que l'existence de phénomènes localement destructeurs ayant pu avoir des impacts importants (PNUD, 2016).

Dans ce contexte, l'évaluation d'un projet d'adaptation de l'agriculture aux changements climatiques doit, par exemple, impérativement analyser les rendements obtenus au regard des conditions climatiques rencontrées au cours de la phase d'exécution, afin de révéler l'efficacité réelle des activités, puisque les conditions climatiques influencent fortement la productivité agricole, quelles que soient les activités d'appui réalisées. Une bonne répartition des pluies peut, en effet, entraîner une augmentation spectaculaire des rendements, sans que ce résultat soit nécessairement imputable aux activités du projet. Inversement, des rendements relativement modestes obtenus au cours d'une saison culturale marquée par des conditions climatiques particulièrement difficiles, peut révéler l'efficacité des techniques et moyens déployés pour limiter les impacts d'une crise climatique et relancer la production en cas de pertes. Toutefois, cette analyse est souvent rendue difficile par le manque de précision des données des agences météorologiques nationales, qui fournissent généralement des moyennes mensuelles à l'échelle régionale, et qui s'avèrent insuffisantes pour retracer les variations saisonnières et les phénomènes intenses ponctuels, tels que les sécheresses localisées ou les fortes précipitations à l'origine de graves inondations, y compris en zone sahélienne. Lorsque des données plus précises existent, c'est leur accès qui reste souvent un défi pour l'évaluateur, malgré l'appui des partenaires locaux.

### **L'évaluation des facteurs de vulnérabilité, un domaine pluridimensionnel et complexe**

L'analyse des conditions climatiques ne suffit pas, à elle seule, à déterminer les facteurs de vulnérabilité aux aléas climatiques. Les impacts des phénomènes climatiques dépendent, en effet, dans une large mesure, d'autres facteurs naturels et anthropiques parmi lesquels : la topographie, le type de sols et la dynamique érosive, le niveau de dégradation de la couverture végétale, la dynamique hydrologique,



l'exposition des terrains et des infrastructures aux zones à risque (inondables, érosives, soumises aux sécheresses, etc.), les conditions d'urbanisation, la qualité des techniques et les matériaux utilisés, le niveau de vie, les modes d'organisation d'une société, etc. (Welborn, 2018). Une bonne évaluation doit donc tenir compte de tous ces facteurs pour déterminer l'état de référence et le niveau de vulnérabilité des composantes biophysiques et socio-économiques, afin de pouvoir évaluer s'ils ont été renforcés et dans quelle mesure (Canadian Council of Ministers of the Environment, 2021). Seule l'analyse de ces facteurs de vulnérabilité permet de réaliser une évaluation objective et rigoureuse du niveau de renforcement des capacités d'adaptation aux changements climatiques, ce qui est complexe et exige une bonne connaissance thématique de la part de la personne évaluatrice.

### **La pertinence relative des indicateurs du cadre de mesure du rendement des projets d'adaptation aux changements climatiques**

Face à la complexité de la thématique, à la difficulté d'accès aux données et à l'incertitude de l'évolution des conditions climatiques, il est fréquent que les organisations responsables de la mise en œuvre des projets définissent des indicateurs de rendement, ou de progression, plus faciles à évaluer que les indicateurs d'impacts ou de résultats, mais qui sont rarement les plus pertinents pour évaluer le renforcement des capacités d'adaptation. Ce constat s'explique, en partie, par le souci de définir des résultats et des cibles facilement atteignables pour garantir la réussite du projet vis-à-vis du bailleur de fonds, et réduire ainsi les risques institutionnels et financiers dans un contexte à la fois changeant, difficilement prévisible et non contrôlable. Les contraintes de temps et de budget incitent aussi les organisations à choisir des indicateurs facilement mesurables, se limitant souvent à la définition la plus restreinte des indicateurs SMART – Spécifiques, Mesurables, Acceptables, Réalistes, Temporellement définis.

Ainsi, de nombreuses interventions proposent, par exemple, des indicateurs tels que la perception des populations cibles quant à l'amélioration de leurs capacités d'adaptation, alors que les personnes interrogées n'ont bien souvent qu'une compréhension limitée de ce concept. D'autres sont basées sur les rendements agricoles bruts sans corrélation avec les conditions climatiques ou la superficie des zones restaurées ou reboisées, sans que leur contribution directe à la conservation des sols ou à la réduction des risques et désastres soit clairement démontrée (Van Gameren *et al.*, 2014). D'une manière générale, le manque d'indicateurs et de critères d'évaluation qualitatifs constitue une limite évidente pour évaluer l'efficacité et l'efficience d'un projet d'adaptation aux changements climatiques.

### **Le temps limité et contraint de l'évaluation**

L'évolution du climat, ses impacts et l'efficacité des mesures d'adaptation mises en place ne s'appréhendent généralement pleinement que sur le long terme. Or, la durée de la plupart des projets de développement dépasse rarement cinq ans, et ces derniers sont généralement évalués après seulement quatre années de mise en œuvre effective (Prowse et Snilstveit, 2010). Ces projets visent, le plus souvent, des résultats « durables », alors que cette durabilité est difficilement démontrable à court terme, puisqu'elle dépend, par définition même, de la capacité des résultats à se maintenir dans le temps, c'est-à-dire une fois l'appui technique et financier du projet arrivé à son terme. De même, les projets d'adaptation visant à conserver ou à restaurer le milieu naturel à travers des activités de protection des sols, de reboisement ou de conservation des ressources hydriques ne produisent des effets significatifs, visibles et mesurables qu'à moyen ou à long terme, ce qui permet difficilement de les évaluer en fin de projet (Prowse et Snilstveit, 2010). La méthodologie d'évaluation des résultats doit

alors s'adapter pour mesurer une réalité complexe et changeante, dont l'évaluation objective et quantifiable exigerait des études scientifiques poussées, menées sur le long terme. Toutefois, les évaluations *ex post* restent rares.

À ce défi s'ajoute celui de la durée et de la saisonnalité de l'évaluation. Notre expérience démontre que la personne évaluatrice dispose d'un temps extrêmement limité sur le terrain, quelques semaines tout au plus, et elle maîtrise rarement le calendrier de la phase de collecte des données, ce qui limite considérablement la capacité d'analyse des conditions biophysiques, et oriente la méthodologie de terrain vers une collecte de données reposant principalement sur des diagnostics rapides et des témoignages. Là encore, l'analyse de ces données exige une bonne maîtrise de la thématique, afin d'identifier les effets réellement attribuables aux activités du projet. Il arrive en effet fréquemment, toujours d'après notre expérience, que les populations cibles et même les organisations partenaires, les autorités locales et les services techniques de l'État attribuent les problèmes de production aux changements climatiques, thématique très médiatisée ces dernières années, sans que le lien direct avec l'évolution des phénomènes météorologiques soit clairement et rigoureusement établi. Les aléas climatiques sont presque systématiquement mis de l'avant pour expliquer une crise, alors que l'influence de l'évolution des facteurs de vulnérabilité (déforestation, artificialisation et épuisement des sols, surexploitation des nappes et des cours d'eau, colonisation des zones à risque, etc.) sur l'ampleur de la catastrophe est largement sous-estimée.

Face aux nombreux défis soulevés pour l'évaluation d'un projet d'adaptation aux changements climatiques dans les pays en développement, plusieurs approches peuvent être développées pour assurer la qualité et la rigueur de la démarche méthodologique.

### **La diversification des sources d'information**

Face au manque de précision des données météorologiques dans la plupart des pays en développement, les données collectées par d'autres institutions ou organisations peuvent fournir un complément d'information souvent indispensable à la compréhension plus fine de l'évolution du climat et de ses impacts sur la disponibilité en eau, la production agricole ou les conditions de vie en zone rurale. Les ministères de l'Agriculture ou les services de sécurité civile s'avèrent, par exemple, une bonne source d'information pour identifier les phénomènes climatiques et estimer leurs impacts et leur influence potentielle sur un projet. En Haïti, par exemple, le ministère de l'Agriculture des Ressources naturelles et du Développement rural (MARNDR) abrite l'Unité HydroMétéorologique d'Haïti (U.H.M.), qui dispose de données de référence sur le climat qu'il est possible de corréliser aux rendements agricoles dans les différents départements du pays, même si la topographie et la qualité des sols influencent aussi fortement ces résultats.

Le manque de précision des données climatiques quantitatives disponibles auprès des services météorologiques nationaux doit aussi inciter la personne évaluatrice à compléter ces informations par des données qualitatives issues de témoignages recueillis et d'enquêtes réalisées auprès des populations et des parties prenantes du projet. Ces informations permettent généralement d'appréhender de façon précise les impacts des phénomènes climatiques sur le territoire d'intervention, les activités du projet et la qualité de vie des populations, et ce, du point de vue des bénéficiaires. L'analyse croisée de ces deux sources d'informations, quantitative et qualitative, permet de mieux comprendre le contexte, ainsi que la contribution réelle des activités du projet sur l'amélioration des conditions de vie et d'adaptation des personnes ciblées par le projet.

## Le recours aux méthodologies scientifiques

Pour démontrer les effets positifs de certaines activités sur la réduction de la vulnérabilité et des impacts des changements climatiques, il est indispensable d'appuyer l'évaluation par des études scientifiques, qui démontrent le lien entre l'activité menée et les effets escomptés. Ainsi, même s'il est difficile de prouver la contribution directe des actions de conservation du milieu naturel sur l'amélioration de la capacité de recharge des nappes, la lutte contre l'érosion et les inondations ou la régulation des microclimats, le lien scientifique établi entre ces diverses composantes permet de considérer ces résultats comme valides, à condition que l'évaluation de terrain démontre que les techniques employées sont adaptées au milieu et au résultat souhaité. Un reboisement efficace pour conserver les sols, pour conserver les ressources en eau et pour préserver la biodiversité doit ainsi être réalisé sur courbe de niveau, c'est-à-dire perpendiculairement à la pente, avec des espaces locaux et diversifiés, qui permettent la reconstitution des écosystèmes forestiers. Les reboisements intensifs réalisés à partir d'eucalyptus ou pour solliciter des crédits carbone ont souvent eu des effets très négatifs sur la conservation des ressources en eau et sur la préservation de la biodiversité. D'autres espèces animales ou végétales invasives allogènes introduites dans le cadre de projets de développement ont eu des effets négatifs importants sur le fonctionnement des écosystèmes locaux.

Puisque les études scientifiques spécifiques ne sont pas toujours disponibles à l'échelle pertinente, dans les régions d'intervention ou pour la thématique souhaitée, les méthodes utilisées en écologie, en géographie ou en géologie et pédologie peuvent également être utilisées pour évaluer le résultat de certaines activités. Des techniques simples d'échantillonnage, telles que les transects ou quadrats, peuvent ainsi permettre d'évaluer facilement et relativement rapidement l'évolution des conditions naturelles et des facteurs de vulnérabilité. Certains projets prévoient la mise en place de stations météorologiques automatisées dans les zones d'intervention, afin d'assurer les suivis des conditions climatiques et orienter les choix productifs<sup>1</sup>. D'autres installent des échelles limnimétriques pour mesurer l'évolution des niveaux d'eau d'un cours d'eau, ou des moulinets courantomètres pour suivre l'évolution des débits.

Mais l'un des outils les mieux adaptés pour réaliser un diagnostic et un suivi environnemental à grande échelle est sans doute la télédétection spatiale, en particulier pour les projets agricoles ou en foresterie, car elle permet de retracer la dynamique de la couverture végétale et l'évolution de l'occupation des sols, selon une perspective historique, c'est-à-dire bien avant le démarrage d'un projet. La création des images, par classification du rayonnement électromagnétique en fonction des longueurs d'onde (spectre électromagnétique), permet de fournir un très grand nombre d'informations sur l'utilisation des sols, en particulier dans l'infrarouge et le proche infrarouge, notamment la composition de la végétation, le type de cultures, leur état de maturité et de santé, les rendements potentiels, le niveau d'humidité ou de stress hydrique, etc. La cartographie et l'analyse spatiale qui découlent de ces informations constituent des outils puissants d'analyse et d'évaluation environnementale, mais également d'aide à la décision pour les gestionnaires de projet et de communication et de coordination avec les autorités et les populations partenaires. L'analyse croisée de ces images, avec des relevés et des enquêtes de terrain, permettent de dresser un portrait précis, dans l'espace et dans le temps, de l'évolution des facteurs de vulnérabilité et des stratégies d'adaptation aux changements climatiques. Si un

1. Nous faisons ici référence au Projet Innovation et mobilisation pour la sécurité alimentaire (IMSA), financé par Affaires mondiales Canada et mis en œuvre par Mission inclusion au Burkina Faso, en Bolivie et au Pérou (2015-2020).

budget spécifique doit être prévu pour mobiliser les données et l'expertise technique nécessaire à leur exploitation, le coût de la télédétection est généralement largement inférieur à celui d'un suivi terrain régulier, en particulier dans les zones difficilement accessibles.

### **L'analyse critique des résultats et des indicateurs de rendement**

La méthodologie d'évaluation d'un projet d'adaptation aux changements climatiques doit s'appuyer sur une analyse approfondie de la pertinence des indicateurs du cadre de mesure du rendement, et doit prendre en considération l'ensemble des facteurs externes ayant facilité ou freiné l'atteinte des résultats (OECD, 2015). Si ce constat s'applique à toute démarche d'évaluation de projet, il s'avère particulièrement important dans le cas des projets d'adaptation aux changements climatiques, pour lesquels le contexte météorologique changeant et ses impacts difficilement prévisibles et maîtrisables sont l'objet même du projet. La démarche d'évaluation doit privilégier les indicateurs d'impact plutôt que les indicateurs de suivi des activités qui restent trop souvent privilégiés face aux risques et aux défis inhérents à ce type de projets. Les résultats de l'évaluation doivent aussi être analysés et relativisés en considérant l'influence, négative ou positive, des conditions climatiques sur le niveau d'atteinte des résultats, ce qui implique une analyse qualitative approfondie. En effet, certains projets n'ayant pas atteint pleinement les cibles définies peuvent malgré tout avoir été particulièrement pertinents et efficaces face à des conditions climatiques difficiles, alors que d'autres présentant des résultats supérieurs aux objectifs peuvent n'avoir eu qu'un impact limité et avoir principalement bénéficié des conditions climatiques favorables durant la phase d'exécution des activités.

Enfin, la démarche d'évaluation de projet doit s'appuyer sur l'étude d'impact environnemental et social (EIES) – aussi appelée Évaluation environnementale stratégique – en particulier son Plan de gestion environnementale et sociale (PGES). Ce plan est exigé par la plupart des gouvernements et des bailleurs de fonds internationaux (Nations Unies, Banque mondiale, OCDE, etc.) pour les projets susceptibles de générer des effets négatifs importants, en particulier ceux comportant des infrastructures d'envergure (routes, barrages, zones industrielles, sites de décharge, etc.) ou visant l'exploitation des ressources naturelles (projets agricoles, d'irrigation, hydroélectriques, forestiers, miniers, etc.). De plus, il fournit des informations importantes sur le contexte biophysique et socio-économique de la zone d'intervention, présente le cadre juridique environnemental du pays d'accueil et identifie les effets potentiels des interventions proposées. Le PGES présente les mesures d'évitement, d'atténuation et de compensation des effets négatifs potentiels ainsi que les mesures de renforcement de ses effets positifs. Il fait également l'objet d'un processus d'évaluation et doit être considéré comme un élément central de toute démarche d'évaluation de projet.

## **Conclusion**

L'évaluation de projets de coopération internationale sur l'adaptation aux changements climatiques est un domaine complexe qui exige une approche multidisciplinaire, le recours à des méthodologies rigoureuses et des outils d'analyse performants et innovants. L'intérêt de plus en plus marqué des bailleurs de fonds pour l'évaluation des impacts environnementaux, la recherche d'efficacité et le contexte de crises multiples de ces dernières années (insécurité, catastrophes naturelles, crise sanitaire, etc.) accélèrent le recours à l'innovation technologique (télédétection spatiale, drones, etc.) et à la délocalisation de l'expertise, qui permettent un suivi de proximité à moindre coût et une bonne compréhension des dynamiques économiques, sociales et culturelles qui façonnent les stratégies d'adaptation aux changements climatiques.

## Pour aller plus loin

- Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC). (2019). *Résumé à l'intention des décideurs, Résumé technique et Foire aux questions*. [https://www.ipcc.ch/site/assets/uploads/sites/2/2019/09/SR15\\_Summary\\_Volume\\_french.pdf](https://www.ipcc.ch/site/assets/uploads/sites/2/2019/09/SR15_Summary_Volume_french.pdf)

Publication de référence permettant de cerner les principaux enjeux en matière de changements climatiques et leurs impacts sur les populations vulnérables à court, moyen et long terme.

- Ministère de l'Environnement et de la Lutte contre les changements climatiques. (2021). *Les changements climatiques et l'évaluation environnementale : Guide à l'intention de l'initiateur de projet*. Gouvernement du Québec. <https://www.environnement.gouv.qc.ca/evaluations/directive-etude-impact/guide-intention-initiateur-projet.pdf>

Décrit de façon pratique le contenu de l'évaluation environnementale des interventions portant sur les changements climatiques.

- Ministère fédéral de la Coopération économique et du Développement (GIZ). (2015). *Suivre et évaluer l'adaptation au changement climatique à haut niveau : analyse comparative de dix systèmes*. [https://www.adaptationcommunity.net/download/me/national-level-me\(2\)/giz2015-fr-suivre-evaluer-l-adapt-change-clima.pdf](https://www.adaptationcommunity.net/download/me/national-level-me(2)/giz2015-fr-suivre-evaluer-l-adapt-change-clima.pdf)

Présente une analyse comparative des différentes approches et méthodologies d'évaluation politiques nationales, en plus de présenter les principaux concepts, définitions et outils utilisés.

## Références bibliographiques

Canadian Council of Ministers of the Environment. (2021). *Guidance on Good Practices in Climate Risk Assessment*. <https://ccme.ca/en/res/riskassessmentguidancesecured.pdf>

IPCC. 2021. *Climate Change 2021 : The Physical Science Basis*. [https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC\\_AR6\\_WGI\\_Full\\_Report.pdf](https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Full_Report.pdf)

OECD. (2015). *National Climate Change Adaptation : Emerging Practices in Monitoring and Evaluation*. [https://read.oecd-ilibrary.org/environment/national-climate-change-adaptation\\_9789264229679-en#page1](https://read.oecd-ilibrary.org/environment/national-climate-change-adaptation_9789264229679-en#page1)

PNUD. (2016). *Une vision nouvelle pour les services météorologiques et climatologiques en Afrique*. [https://www.undp.org/sites/g/files/zskgke326/files/publications/WeatherAndClimateServicesAfrica\\_FR.pdf](https://www.undp.org/sites/g/files/zskgke326/files/publications/WeatherAndClimateServicesAfrica_FR.pdf)

Prowse, M. et Snilsveit, B. (2010). *Impact Evaluation and Interventions to Address Climate Change: A Scoping Study*. *Journal of development Effectiveness*, 2(2). [https://www.researchgate.net/publication/46448679\\_Impact\\_Evaluation\\_and\\_Interventions\\_to\\_Address\\_Climate\\_Change\\_A\\_scoping\\_study](https://www.researchgate.net/publication/46448679_Impact_Evaluation_and_Interventions_to_Address_Climate_Change_A_scoping_study)

Van Gameren, V., Weikmans, R. et Zaccai, E. (2014). *Mesures et suivi de l'adaptation*. Dans V. Van Gameren, R. Weikmans, et E. Zaccai (dir.), *L'adaptation au changement climatique* (p. 91-102).

Welborn, L. (2018). *Africa and climate change : Projecting vulnerability and adaptive capacity* (publication n° 14). Institute for Security Studies. [https://media.africaportal.org/documents/Africa\\_and\\_climate\\_change.pdf](https://media.africaportal.org/documents/Africa_and_climate_change.pdf)



# Genèse et perspectives d'un mouvement africain de l'évaluation

*Oumoul Khayri Ba Tall*

*Carla-Anide Guillaume*

*Vénétia Sauvain*

## Introduction

*Made In Africa Evaluation (MAE)*, ou modèle d'évaluation adaptée à l'Afrique (EAA) en français, est un mouvement qui appelle à une École Africaine d'Évaluation. Il a été initié par l'Association Africaine d'Évaluation (AfrEA) en marge de la conférence de Niamey au Niger en 2007, par un groupe d'évaluateurs et de chercheurs africains conviés par la professeure Zenda Ofir de l'Afrique du Sud, pour réfléchir sur la façon de concevoir une évaluation plus conforme aux différents contextes africains et à la réalité de ce continent. L'appel a suscité beaucoup d'enthousiasme, soulevé de grandes attentes, mais a donné lieu à un nombre limité d'études approfondies et de publications sur la question, à ce jour. Une étude menée en 2015, par la professeure Bagele Chilisa de l'Université du Botswana pour le compte de l'AfrEA, a eu le mérite de revisiter le concept dans son processus historique, de poser et de clarifier les hypothèses épistémologiques, de proposer quelques orientations et pistes de réflexion avant d'appeler à une plus grande contribution des auteurs africains au développement des théories et pratiques de l'évaluation d'obédience africaine.

Ce chapitre propose d'explorer la raison d'être du mouvement MAE ainsi qu'un certain nombre d'enjeux qu'il soulève, notamment pour sa mise en œuvre en évaluation.

## Les contours du concept *Made in Africa*

Sur le continent africain, l'évaluation est influencée par les agences d'aide au développement dont les modèles et pratiques sont véhiculés à travers les exigences des accords de l'aide au développement, qui encadrent le travail des praticiens de l'évaluation et des consultants des pays donateurs ou des agences concernées (Cloete, 2016, p. 57). Les bailleurs de fonds favorisent des interventions basées sur le concept de déficit («manque de quelque chose») pour relever les défis de l'Afrique et pour formuler les stratégies d'évaluation, leurs méthodologies et leurs processus de diffusion (Chilisa *et al.*, 2017, p. 315). Or, les approches de résolution des problèmes, les méthodes d'évaluations et leur processus de diffusion, plutôt issus des pays du Nord, ont de la difficulté à s'incarner dans les contextes africains. Le mouvement MAE pose la question : si l'évaluation était d'obédience africaine, à quoi ressemblerait-elle?



## **Une rhétorique pour la décolonisation effective et l'affirmation de l'Afrique**

En réalité, la rhétorique de *Made in Africa* s'inscrit dans un mouvement plus global et transversal à travers différentes sphères d'activités, disciplines et écoles de pensée. Elle appelle à une relecture des relations entre l'Afrique et le reste du monde, pour les débarrasser des attitudes paternalistes et néocoloniales et les rendre plus équilibrées sous le prisme d'une plus grande équité, d'une meilleure justice sociale entre les nations et les peuples, de la considération et d'un respect réciproque. Au-delà de la nécessité historique de rétablir les fondements d'une coopération internationale, l'échec des modèles de développement postcoloniaux, combiné aux crises successives multiformes, ont amené les acteurs du renouveau à une prise de conscience progressive, soit qu'il n'existe pas de solution durable aux problèmes de développement de l'Afrique en dehors d'une approche enracinée et ancrée sur les modes de pensée, les modes d'action et « les façons d'Être » de l'Africain en harmonie avec les contextes des sociétés africaines. L'« africanisation de l'évaluation » renvoie à un processus qui place la philosophie africaine, les visions du monde, les systèmes de connaissances et les valeurs africaines au centre de l'évaluation (Chilisa, 2012, p. 12). Tout en reconnaissant la formidable diversité des sociétés, des us et coutumes qui existent en Afrique et à travers tous les peuples du continent, les socioanthropologues et les historiens ont mis en exergue un grand nombre de similitudes et de convergences qui justifient une approche fédératrice qui serait valable à travers le continent : dans les façons d'Être, l'organisation de la famille et de la cité, etc. Un certain nombre de dimensions qui seraient caractéristiques de la « Conscience Africaine » : les liens entre les Êtres et les Choses, entre la Nature et la Société, entre les Générations, à travers la Vie et la Mort, etc. Pour Bagele, un dénominateur commun indéniable à travers les sociétés africaines est la notion de UBUNTU, dont il existe une dénomination dans bon nombre de langues africaines, qui se traduit littéralement par « je suis parce que tu es ». Cette philosophie valorise le partage et l'appropriation collective des opportunités, des responsabilités et des défis, l'importance des personnes et des relations en rapport avec les choses matérielles, la prise de décision participative et le leadership, la réconciliation comme objectif de gestion et de résolution des conflits, entre autres (Cloete, 2016, p. 63). Les méthodes d'évaluation qui mettent en évidence les expressions quotidiennes du bien-être et l'expérience vécue, par la culture ou le langage, entre autres, sont particulièrement appropriées dans ce contexte.

## **Un contexte international favorable à l'évaluation pour un développement endogène ancré dans les valeurs locales**

Plusieurs évolutions du contexte international ont permis l'émergence du mouvement MAE. D'abord, l'Agenda 2030 et l'adoption des 17 objectifs de développement durable (ODD), selon une architecture de la coopération internationale qui met en avant la responsabilité collective des pays dans l'atteinte des ODD, et leur évaluation. Ensuite, la promotion du renforcement des capacités en évaluation, y compris dans la mise en place des systèmes nationaux d'évaluation (2015 déclarée par les Nations Unies année internationale de l'évaluation). Enfin, la complexité croissante des problèmes globaux (pandémies, changements climatiques, conflits et terrorisme, migrations, etc.) a mis les gestionnaires publics devant le défi de disposer de données probantes et à temps opportun, pour être en mesure de prendre des décisions idoines; ceci dans un contexte d'une demande de redevabilité et de transparence de plus en plus forte. Les méthodes et approches d'évaluation classiques évoluent pour contribuer à la réponse à ces problèmes interconnectés, dont les conséquences sont particulièrement prégnantes au niveau local (par exemple, une inondation, un massacre dans un village, un naufrage de barque de migrants). Par ailleurs, l'adoption d'une vision africaine du développement avec l'Agenda 2063 est venue confirmer la détermination affichée par l'Afrique de prendre son destin en main en affirmant son identité propre.



Dans ce contexte, la thématique MAE s'est invitée progressivement dans le débat en tant que réponse africaine à la problématique de l'évaluation du développement, ce que Adeline Sibanda, ancienne présidente de l'AfrEA, a résumé en ces termes dans le journal de la Banque Africaine de Développement (BAD) consacré au sujet : « L'objectif général de l'EAA est d'accroître l'influence des méthodes, théories et philosophies ancrées en l'Afrique dans le développement et les évaluations à l'échelle mondiale, particulièrement celles qui sont censées profiter à l'Afrique. » (BAD/IDEV, 2019, p. 11).

## Enjeux de la mise en œuvre du mouvement MAE

### Enjeu 1 : articuler le concept par un contenu substantiel

Le concept « MAE » nécessite d'être clarifié en lui donnant un contenu concret qui justifie sa pertinence et son utilité.

L'idée « MAE » lancée à Niamey et développée par Bagele (2015) a été conceptualisée par un groupe de 19 « leaders éclairés/penseurs » africains (*African Thought Leaders* en anglais), réunis à Bellagio (Italie) du 14 au 17 novembre 2012. Dans le rapport de Bellagio, considéré par certains comme la première contribution substantielle à la clarification du concept (Cloete, 2016), le MAE est présenté comme une évaluation impulsée par et enracinée en Afrique. Malgré cette avancée sur l'articulation du concept, le rapport reconnaît qu'il reste beaucoup à faire en lançant un appel à des recherches complémentaires pour apporter plus de substance.

Le premier effort de définition est proposé dans un article paru récemment dans une édition du *Journal africain de l'évaluation* (Omosa *et al.*, 2021). Les auteurs définissent le MAE comme une « évaluation conduite sur la base des standards de l'AfrEA; qui utilise des méthodes et approches contextualisées, dans le but d'aligner toutes les évaluations aux styles de vie et besoins des populations africaines ciblées, tout en assurant la promotion des valeurs africaines » (Omosa *et al.*, 2021, p. 6)<sup>1</sup>. Cette définition est certes claire dans son énoncé, mais il faudra attendre de voir les standards africains (rebaptisés « principes » dans la nouvelle version à venir) à l'œuvre pour être en mesure d'en apprécier l'opérationnalité.

### Enjeu 2 : Sauvegarder la crédibilité et l'universalité de l'évaluation en assurant sa pertinence pour l'Afrique

La question qui se pose depuis le début de ce mouvement est de savoir comment faire en sorte que l'évaluation d'obédience africaine puisse s'affirmer dans son originalité tout en conservant le caractère scientifique qui fonde sa crédibilité et son universalité. Il s'agit alors de sauvegarder les caractéristiques d'une évaluation de qualité sur les plans technique et scientifique, tout en lui assurant une plus grande utilité pour orienter et soutenir un développement harmonieux qui serve les intérêts des populations et des acteurs endogènes en priorité.

---

1. Comme il s'agit d'une traduction libre des auteures, voici la version en anglais pour éviter des malentendus : « *Evaluation that is conducted based on AfrEA standards, using localised methods or approaches with the aim of aligning all evaluations to the lifestyles and needs of affected African peoples while also promoting African values* ».

D'aucuns disent que l'évaluation en tant que méthodologie devrait être conçue comme une discipline qui transcende les particularités sociétales, à la fois temporelles et spatiales, et ne saurait être qualifiée d'une quelconque délimitation sans perdre son caractère « scientifique ». Le mouvement MAE reconnaît justement cette universalité de l'évaluation comme approche d'analyse de l'action publique, tout en rappelant la nécessité de faire usage des « outils et grilles de lecture » appropriés aux différents contextes. Dans l'état actuel du développement de la théorie et des pratiques, cette particularité est perçue comme une forme d'application des approches participatives au contexte africain, ce qui revient à considérer le MAE comme un « principe » ; ce qui diminue sa portée méthodologique potentielle.

### Enjeu 3 : Méthodologie et boîtes à outils

À travers la métaphore africaine de l'arbre d'évaluation, Chilisa *et al.* (2017) propose une conception globale du MAE qui va du plus faible niveau d'indigénisation (où l'adaptation consiste en une plate traduction des concepts et outils), aux niveaux les plus achevés où l'évaluation est réellement enracinée dans la vision africaine du monde. L'auteure introduit ainsi le paradigme africain d'évaluation relationnelle comme pouvant servir de fondement à des pratiques et théories de l'évaluation enracinées dans une vision du monde africaine. Différentes indications pratiques proposent ce que devrait être une approche MAE, ou au contraire ce qu'il convient d'éviter.

Il faut éviter de se limiter à une adaptation accessoire et à la simple traduction des instruments d'évaluation dans les langues locales (Chilisa *et al.*, 2017, p. 325). L'adaptation des modèles occidentaux, des théories et des instruments doit permettre de les rendre contextuellement pertinents, culturellement appropriés et inclusifs des parties prenantes locales et des évaluateurs africains. L'évaluation MAE est fondée sur le savoir et les approches locales, la prise en compte du mode de vie des populations et des traditions africaines (Omosa *et al.*, 2021, p. 8). Les questions d'évaluation ne doivent pas être prédéterminées, mais plutôt élaborées en consultation avec les participants ou la communauté (Chilisa *et al.*, 2017, p. 325). Les données sont analysées avec la communauté ou avec les personnes qui comprennent et qui peuvent interpréter la langue, les expressions, les proverbes et les croyances spirituelles de la communauté locale. Le MAE reflète la logique africaine de la circularité, par opposition à la logique linéaire des méthodes d'évaluations occidentales (Chilisa *et al.*, 2017, p. 323). La réalité africaine doit définir les critères ou normes, les indicateurs de succès ou d'échec des projets et les conclusions sur la valeur ou le mérite des programmes, politiques ou projets de développement (Chilisa *et al.*, 2017, p. 318). Pour que l'évaluation soit enracinée en Afrique, elle doit aussi inclure une analyse de la contribution de l'intervention au bien-être de la communauté, et équilibrer les priorités et indicateurs occidentaux et africains (Chilisa *et al.*, 2017, p. 324).

Cela dit, il y a encore peu de clarté sur une méthodologie commune et un cadre conceptuel fédérateur à l'usage des évaluateurs. La BAD a consacré une édition de sa revue périodique *eVALUation Matters* à des exemples de pratiques de la MAE, dont certaines sont des adaptations de démarches connues aux réalités et contextes locaux : c'est le cas de la méthode Baraza, qui est une adaptation du contrôle citoyen de l'action publique en Ouganda, ou de l'analyse contextuelle, pour être en mesure d'identifier les bonnes questions, ou de déterminer l'évaluabilité au préalable. D'autres cas illustrent comment les outils classiques n'ont pas permis de détecter les effets induits par une intervention, par exemple lorsque les données n'existent pas sous les formes connues (documentées, accessibles), ou que le cadre

de l'évaluation n'est pas propice (temps, langage, lieu, etc.). Les auteurs estiment aussi que les défis et les échecs des évaluations en Afrique illustrent encore plus l'inadéquation des méthodes classiques, et préconisent que leur exploration soit favorisée, pour améliorer les façons de faire. Les promoteurs du MAE sont conscients de la nécessité de continuer à développer la théorie et les modèles d'évaluation, et encouragent la recherche qui met en avant les réalités africaines et les cultures locales.

#### **Enjeu 4 : Sphère d'influence et capacité à utiliser le MAE**

Le succès du MAE sera déterminé par son degré d'adoption dans les évaluations, surtout celles qui concernent l'Afrique. Il y a deux conditions qui vont favoriser ce développement : l'adhésion franche des acteurs du côté de la demande comme du côté de l'offre, mais aussi leur capacité à acquérir les connaissances et aptitudes nécessaires à la mise en œuvre du MAE.

Dans la communauté des évaluateurs en Afrique, l'engouement affiché en faveur du MAE est visible et audible, et cela même avant d'en voir une compréhension partagée. Il s'agit beaucoup plus d'une adhésion dans les principes fondateurs, mais il en faudra bien plus pour arriver à la mise en pratique. Parmi les acteurs de développement en Afrique, il y a la communauté internationale qui joue encore un rôle non négligeable, voire déterminant dans certains contextes. Aujourd'hui encore, l'essentiel des budgets d'évaluation provient de ressources extérieures, de la coopération multilatérale ou bilatérale. Reste à savoir dans quelle mesure le MAE sera compris et promu dans les pratiques des partenaires du Nord, surtout lorsque l'évaluation portera sur des actions qui les impliquent.

## **Conclusion**

Le MAE se présente comme l'expression d'une « École Africaine d'Évaluation » naissante, en phase avec les aspirations du moment, avec la décolonisation effective et l'affirmation de l'Afrique, pour enfin assurer son développement effectif au profit de ses populations. L'idée n'est pas qu'elle s'isole du reste du monde, mais bien qu'elle contribue à l'enrichissement de la théorie, de la pratique et de la culture évaluative dans une dynamique d'enracinement et d'ouverture, notamment avec les apports d'autres régions du monde (paradigme asiatique d'évaluation (Russon, 2008), approche théorique Kaupa Maori (Kerr, 2012)), aux côtés des approches, méthodes et outils classiques du Nord qui font encore leurs preuves. L'AfrEA a, du reste, intégré le concept (avec la révision en cours des Principes Directeurs de l'Évaluation en Afrique) dans sa boîte à outils en reconnaissant et en appuyant l'intérêt grandissant pour « une École Africaine d'Évaluation, adaptée par et pour les contextes africains, les besoins et les systèmes de connaissances, tout en s'inspirant des théories et bonnes pratiques internationales » (AfrEA, s.d.).

## Références bibliographiques

- AfrEA. (s.d.). [Boîte à outils sur les principes africains de l'évaluation à paraître dans le cadre de la 10<sup>e</sup> conférence de l'AfrEA, Addis Ababa]
- Chilisa, B. (2012). Made in Africa Evaluation : Uncovering African Roots in Evaluation Theory and Practice. Dans The Bellagio Centre – The Rockefeller Foundation (dir.), *African Thought Leaders Forum on Evaluation and Development*. Centers for Learning on Evaluation and Results, AfrEA et The Graduate School of Public and Development Management.
- Chilisa, B., Major, T. E., Gaotlhobogwe, M. et Mokgolodi, H. (2017). Decolonizing and indigenizing evaluation practice in Africa : Toward African relational evaluation approaches. *The Canadian Journal of Program Evaluation*, 30(3), 312–328. <https://evaluationcanada.ca/system/files/cjpe-entries/30-3-313.pdf>
- Cloete, F. (2016). Developing an Africa-rooted programme evaluation approach. *African Journal of Public Affairs*, 9(4). <http://hdl.handle.net/2263/59022>
- Évaluation indépendante du développement de la Banque Africaine de Développement (BAD/IDEV). (2019). Des évaluations adaptées à l'Afrique. Volume 1 : Approches théoriques. *eVALUation Matters*.
- Évaluation indépendante du développement de la Banque Africaine de Développement (BAD/IDEV). (2019). Des évaluations adaptées à l'Afrique. Volume 2 : Applications pratiques. *eVALUation Matters*.
- Kerr, S. (2012). Kaupapa Māori Theory-based Evaluation. *Evaluation Journal of Australasia*, 12(1), 6-18. <https://doi.org/10.1177/1035719X1201200102>
- Omosa, O., Archibald, T., Niewolny, K., Stephenson, M. et Anderson, J. (2021). Towards defining and advancing 'Made in Africa Evaluation'. *African Evaluation Journal*, 9(1). <https://doi.org/10.4102/aej.v9i1.564>
- Russon, C. (2008). An Eastern Paradigm of Evaluation. *Journal of Multidisciplinary Evaluation*, 5(10), 71-77. [https://journals.sfu.ca/jmde/index.php/jmde\\_1/article/download/183/203](https://journals.sfu.ca/jmde/index.php/jmde_1/article/download/183/203)

# L'évaluation des politiques publiques en Afrique : pratiques, contraintes et perspectives

*Bobo B. Kabungu*

## Introduction

L'évaluation des politiques publiques a déjà pris son envol en Afrique. Mais le parcours reste long. Après avoir fait le lien entre la qualité des institutions et l'évaluation des politiques publiques (section 1), ce chapitre présentera brièvement la pratique évaluative sur le continent (section 2), avec un intérêt sur les contraintes observées sur le terrain et les perspectives d'institutionnalisation (section 3).

## Qualité des institutions et évaluation des politiques publiques

L'impact de la qualité des institutions sur le développement n'est plus à prouver. Les premières recherches les plus fouillées ont été lancées dans les années 1990 (Levine et Renelt, 1992) avant de s'intensifier dans la décennie 2000 (Acemoglu *et al.*, 2003; Hausmann *et al.*, 2004). Abordant la question à l'aune des réalités africaines en général et congolaises en particulier, Matata Ponyo souligne que « les différences de niveau de vie constatées à l'échelle mondiale s'expliqueraient par la qualité des institutions » (Matata Ponyo, 2016, p. 27). Le rôle des institutions est donc déterminant, d'une part pour stimuler la croissance et lui conférer une durabilité et, d'autre part, pour favoriser et encadrer les transformations structurelles et sociales indispensables au développement. Comme le dit Nangia, « évaluer l'impact grâce à des méthodes scientifiques rigoureuses demeure essentiel pour que le processus de développement repose sur des fondations solides et saines » (Nangia, 2015, p. 4).

Ces institutions peuvent être *hard*, à l'exemple des structures et/ou organes officiels chargés des services principalement étatiques ou *soft*, à l'instar des comportements existants et/ou des mentalités inscrites dans la culture de différentes sociétés. Parmi ces institutions dites *soft*, il sied de citer la culture de la ponctualité, le civisme fiscal (paiement de l'impôt), l'éveil patriotique en contexte de guerre, la promotion de la recherche, la transparence, la pratique évaluative, etc.

En Afrique, si les institutions étatiques commencent timidement à prendre le dessus sur leurs animateurs dans un contexte de démocratisation progressive, le constat est plus sombre en matière d'appropriation de l'évaluation (des politiques publiques). En effet, en dehors d'une pression liée à un éventuel renouvellement de financement d'un programme par un bailleur de fonds extérieur, l'évaluation est soit reléguée, soit bâclée. En présentant sommairement la pratique évaluative en Afrique, ses contraintes et ses perspectives en termes d'institutionnalisation, ce sous-chapitre veut dresser un tableau de cette question, vu de l'intérieur, avec un regard critique.

## La pratique évaluative en Afrique

L'évaluation d'une politique publique ou d'un programme consiste, en général, en une appréciation de l'intervention à l'étude à l'aune de sa pertinence, de son efficacité, de son efficience, de son impact et de sa viabilité. En ce sens, elle se rapproche de la qualité publique, et sa mise en pratique devient l'un des soubassements de la bonne gouvernance, car tout gestionnaire de service public s'attend à un jugement sur sa capacité à atteindre les résultats escomptés et son « aptitude à satisfaire les besoins des usagers, besoins exprimés ou implicites » (Beau *et al.*, 2005, p. 11).

Considérée comme telle, l'évaluation en Afrique demeure à un stade rudimentaire, particulièrement au sud du Sahara, malgré quelques cas encore isolés, mais très encourageants (Afrique du Sud, Maroc, Bénin, Rwanda...), et ce, près de 60 ans après ses premiers pas en Amérique du Nord. Du côté de l'Afrique du Sud, l'on note que le pays s'est engagé, depuis la fin de l'apartheid, dans des mesures réformatrices et favorables aux pauvres dont la mise en œuvre nécessite un suivi, une évaluation et une supervision renforcés afin d'apprécier les progrès réalisés (Naidoo, 2011).

Concernant le Maroc, on observe depuis 2001 un regain d'intérêt pour l'évaluation, si l'on s'en tient aux discours du Trône prononcés le 30 juillet de chaque année. En effet, partant d'une analyse textuelle, Lechheb *et al.* (2018) citent les années 2001, 2005, 2006, 2007 et 2008, comme étant celles où l'importance de l'évaluation dans les allocutions des plus hautes autorités du pays atteint son apogée, soulignant la nécessité de créer des organismes en charge de porter un jugement de valeur sur les politiques mises en place dans le Royaume. Par la suite, la Constitution de 2011 a permis l'entrée en fonction de plusieurs instances, notamment le Conseil Économique, Social et Environnemental, la Cour des comptes ainsi que le ministère chargé des Affaires Générales et de la Gouvernance, auxquels des missions respectives de consultation, d'évaluation de projets publics et de vérification de la cohérence des programmes ont été confiées.

Pour sa part, le Bénin sort du lot grâce à la mise en place d'un cadre institutionnel et organisationnel (structure centrale, animateurs compétents et cadre légal et réglementaire) inclusif en matière d'évaluation et qui devrait constituer un substrat pour l'émergence d'une pratique ancrée dans les habitudes des acteurs de la gestion de la chose publique à tous les niveaux. Il est l'un des premiers pays qui s'est doté d'une politique nationale de l'évaluation pour améliorer la gouvernance de l'État, pour apprécier la valeur des interventions publiques et pour optimiser l'utilisation des ressources collectées auprès des contribuables. Cette politique a été adoptée en 2012 « dans le contexte d'accélération de la réforme de gestion axée sur les résultats » (Bureau de l'Évaluation des Politiques Publiques et de l'Analyse de l'Action Gouvernementale, 2019, p. 14), et le Bureau de l'Évaluation des Politiques Publiques et de l'Analyse de l'Action Gouvernementale (BEEPPAAG) a reçu la mission d'en assurer la mise en œuvre, sous la coordination stratégique du Conseil National de l'Évaluation (CNE). Depuis lors, la moyenne annuelle d'exercices d'évaluation a augmenté dans différents secteurs de la vie publique.

Quant au Rwanda, l'évaluation des politiques et des institutions en Afrique (CPIA) classe le pays en tête du classement de l'Afrique subsaharienne, pour l'année 2019, avec une note de 4 sur une échelle de 1 à 6, la moyenne pour la région étant de 3,1. Le Bénin et le Ghana se classent également en ordre utile avec un score ayant progressé de 3,5 à 3,6. La Banque mondiale fait observer que les pays qui se distinguent dans cette notation ont les économies les plus florissantes du continent.

Ce tableau ne pourrait être complet si l'on manquait d'évoquer des exemples beaucoup moins reluisants. La 2<sup>e</sup> Conférence Internationale sur les capacités nationales d'évaluation tenue à Johannesburg en 2011 mettait déjà en exergue une très faible pratique évaluative sur le continent, avec des cas



explicites en Ouganda (Byamugisha et Smith, 2011) et au Malawi (Mangani, 2011) pour ne parler que de ces deux pays, surtout en matière de développement et de lutte contre la pauvreté. À la même période, un examen des capacités, effectué en RDC, en matière de suivi et d'évaluation, avait abouti à un score de 1,1, sur une échelle de 0 à 4. Sept ans plus tard, Kabungu (2018, p. 16) fait observer, à partir de l'analyse textuelle de discours-programme, que l'Exécutif national, « probablement préoccupé par des questions d'ordre existentiel, ne prend pas suffisamment à cœur la question d'évaluation des politiques publiques » (Kabungu, 2018, p. 16).

À ce jour, les autorités publiques dans de nombreux pays d'Afrique continuent de mettre en œuvre de vastes programmes « à impact immédiat » sans qu'une forte amélioration des conditions de vie des bénéficiaires ne s'ensuive. Les grands projets d'infrastructures de transport, d'énergie ou d'habitat lancés çà et là, sans qu'une étude sérieuse n'ait été diligentée au départ, ne serait-ce que pour apprécier les avantages et les coûts, rappellent la série des « éléphants blancs » d'après les indépendances. Le caractère fort budgétivore de ces projets est décrié par le cadre d'évaluation de la gestion des finances publiques (PEFA)<sup>1</sup>. Cet empressement à l'action, parfois sans suffisamment de recul, jette le pont vers les principales contraintes auxquelles se bute l'évaluation en contexte africain.

## Les contraintes de l'évaluation en contexte africain et perspectives

Grâce à une revue documentaire concordante et rapprochée du vécu, les insuffisances ci-après peuvent être citées et classées selon qu'elles se rapportent :

### *À l'offre :*

- L'insuffisance de l'expertise nationale du fait, principalement, du faible intérêt de ces questions, le secteur de l'évaluation restant l'apanage des « experts internationaux » recrutés par les bailleurs de fonds. À ce propos, les filières académiques dans le domaine sont encore rares;

### *À la demande :*

- L'absence de volonté politique, les décideurs craignant probablement que les résultats de l'évaluation ne jouent en leur défaveur. Nombreux sont ceux qui hésitent à financer un exercice dont les conclusions peuvent contredire la pertinence d'une action présentée comme étant la solution la plus efficace à un problème donné. Les décideurs politiques préfèrent les évaluations « qui leur donnent raison », qui confortent leur position et peuvent les aider à gagner de prochaines élections;

1. PEFA : *Public Expenditure and Financial Accountability*. Il s'agit d'un cadre d'appréciation des points forts et des faiblesses de la gestion des finances publiques en recourant à 31 indicateurs quantitatifs réunis autour de sept piliers (fiabilité du budget, transparence des finances publiques, gestion des actifs et des passifs, stratégie budgétaire et établissement du budget fondés sur les politiques publiques, prévisibilité et contrôle de l'exécution du budget, comptabilité et *reporting*, supervision et audit externes) pouvant appréhender la performance de celle-ci via la vérification de l'atteinte des résultats financiers et budgétaires. Mis en place avec le concours de sept partenaires internationaux du développement (la Commission européenne, le Fonds monétaire international, la Banque mondiale ainsi que les gouvernements britannique, français, norvégien et suisse), le programme est devenu, depuis, un outil de référence qui harmonise les points de vue de différentes entités partenaires.

- L'absence d'un cadre institutionnel dûment investi, l'évaluation des politiques publiques n'étant pas formellement dévolue à une entité bien identifiée. Et, dans la mesure où elle l'est, les objectifs et les moyens d'action ne sont pas précisés, rendant ainsi mal-aisée la réalisation de l'activité;
- Le manque de confiance des décideurs dans la démarche : les gestionnaires de programmes préfèrent généralement des évaluations qui appuient les orientations qu'ils avaient prises et ont tendance à rejeter les conclusions de toute évaluation qui met en exergue des insuffisances dans leur action;
- Le déficit d'une planification nationale orientée vers les résultats de développement et l'habitude de la « navigation à vue », plus propice à des évaluations partielles (simple comptabilité des réalisations, sans évoquer la charge qu'elles ont engendrée, ni les prévisions, ni l'impact projeté);
- La négligence de la mesure de la satisfaction des populations cibles des interventions et des services, sauf à l'approche d'une campagne électorale où la qualité publique est effleurée dans les discours;
- Les difficultés pour mettre en œuvre les recommandations : certaines suggestions peuvent aller dans le sens de modifier un schéma prédéfini, ce qui peut sembler lourd, surtout dans le contexte d'une contrainte budgétaire;
- La faible implication de la société civile traduite par le peu ou le manque de demandes de reddition de comptes, en dehors des bailleurs des fonds;
- La non-soutenabilité du coût de l'exercice régulier de l'évaluation des politiques publiques dans un contexte d'amenuisement des recettes publiques, les acteurs préférant investir à autre chose les frais destinés à l'évaluation;
- La méconnaissance des vertus de l'évaluation en termes d'aide à la décision et la fréquente confusion entre l'évaluation, le contrôle, l'audit, le suivi, etc.

***À l'environnement :***

- L'inadéquation du système d'information au besoin d'aide à la décision;
- Une contre-performance de l'administration;
- Un défaut de priorisation des projets suivant leur degré d'importance et de sensibilité;
- L'indisponibilité des statistiques qui rend mal-aisée une évaluation quantifiée;
- Le défaut de considération du temps comme une ressource à part entière : plusieurs réformes sont lancées sans instruments de suivi ni délai de mise en œuvre. On dirait que le temps n'est pas précieux. Les échéances sont souvent placées « pour la forme » avec une possibilité d'être repoussées de façon perpétuelle, autant que les mandats politiques.



Certes, la pratique évaluative est en plein essor en Afrique, avec de plus en plus de professionnels du domaine et une mobilisation autour de la question. Mais les efforts à déployer ne devraient pas s'estomper au regard des défis qui demeurent à relever. Avant tout, il s'agira de « rendre l'évaluation utile et d'instaurer une culture et des normes pour la pérennité du système national d'évaluation » (Djidjoho et Agbota, 2011, p. 98). Et c'est progressivement que l'on s'appropriera l'évaluation d'impact, « un outil essentiel pour mesurer la valeur ajoutée du programme en comparant différents scénarios » (Fule, 2015, p. 102).

## Conclusion

Les contraintes répertoriées ci-dessus ne devraient pas plonger les évaluateurs émergents dans le découragement. Les avancées enregistrées par certains pays constituent une lueur d'espoir. Mais, comme le dit Duran, « le développement de l'évaluation est tributaire [...] de la nature des élites susceptibles de la porter » (Duran, 1993, p. 137). Le rôle de la société civile en général, et du monde académique en particulier, sera donc déterminant dans les années à venir.

## Références bibliographiques

- Acemoglu, D., Robinson, J. A., Thaicharoen, Y. et Johnson, S. (2003). Chiefs : elite control of civil society and economic development in Sierra Leone, *NBER Working Paper* 18691.
- Beau, J., Blanche, F., Garcin, M. et Morice F. (2005). *Démarche qualité et satisfaction du public* [mémoire de recherche, diplôme de Conservateur de Bibliothèque]. <https://www.enssib.fr/bibliotheque-numerique/documents/702-demarche-qualite-et-satisfaction-du-public.pdf?telecharger=1>
- Bureau de l'Évaluation des Politiques Publiques et de l'Analyse de l'Action Gouvernementale. (2019). *Évaluation de la mise en œuvre de la Politique Nationale d'Évaluation (PNE) 2012-2021 et élaboration de son plan d'action 2020-2021*. Secrétariat Général de la Présidence de la République du Bénin.
- Byamugisha, A. et Smith, D. R. (2011). Ouganda : émettre les directives nationales grâce à l'évaluation. Évaluation du Plan d'action de l'Ouganda pour l'éradication de la pauvreté (1997-2007). Dans PNUD, *Actes de la deuxième conférence internationale sur les capacités nationales d'évaluation. Utilisation de l'évaluation dans la prise de décision pour les politiques et les programmes publics* (p. 69-76). Bureau de l'évaluation du PNUD. <http://web.undp.org/evaluation/documents/NEC/2011/NEC-2011Proceedings-FR.pdf>
- Djidjoho, A. N. et Agbota, A. (2011). Bénin : Pratique et utilisation de l'évaluation des politiques publiques. Dans PNUD, *Actes de la 2<sup>ème</sup> Conférence Internationale sur les capacités nationales d'évaluation. Utilisation de l'évaluation dans la prise de décision pour les politiques et les programmes publics* (98-105). Bureau de l'évaluation du PNUD. <http://web.undp.org/evaluation/documents/NEC/2011/NEC-2011Proceedings-FR.pdf>
- Duran, P. (1993). Les ambiguïtés politiques de l'évaluation. *Pouvoirs*, 67, 137-145.
- Fule, C. B. (2015). Une évaluation d'impact rigoureuse pour des résultats de développement accrus et rapides. Dans Évaluation Indépendante du Développement (dir.), *Solutions émergentes aux défis du développement Vol. 2. eVALUatiOn Matters*.

- Hausmann, R., Pritchett, L. et Rodrik, D. (2004). Growth accelerations. *NBER Working Paper 10566*. Cambridge : National Bureau of Economic Research.
- Kabungu, B. B. (2018). La pratique évaluative dans la sphère publique française et congolaise : analyse comparative et perspectives d'institutionnalisation. *Annales de l'Unigom*, 8(2), 1-25.
- Lechheb, H., Ouakil, H. et Jouilil, Y. (2018). Évaluation des politiques publiques : Outil efficace au service du développement. *Revue du Contrôle de la Comptabilité et de l'Audit*, 6, 897-910.
- Levine, R. et Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review*, 82(4), 942-963.
- Mangani, R. (2011). Malawi : Utilisation de l'évaluation de la gestion des résultats de développement. Preuves du Malawi. Dans PNUD, *Actes de la deuxième Conférence Internationale sur les capacités nationales d'évaluation. Utilisation de l'évaluation dans la prise de décision pour les politiques et les programmes publics (121-129)*. Bureau de l'évaluation du PNUD. <http://web.undp.org/evaluation/documents/NEC/2011/NEC-2011Proceedings-FR.pdf>
- Matata Ponyo, M. (2016). *Qualité des institutions et résultats économiques en RDC 1980 – 2015*. L'Harmattan.
- Naidoo, I. A. (2011). La question de l'utilisation des évaluations : Exemples et enseignements tirés de la Commission de la Fonction Publique en Afrique du Sud. Dans PNUD, *Actes de la deuxième Conférence Internationale sur les capacités nationales d'évaluation. Utilisation de l'évaluation dans la prise de décision pour les politiques et les programmes publics (34-41)*. Bureau de l'évaluation du PNUD.
- Nangia, R. (2015). Questions émergentes en matière d'évaluation. Dans Evaluation Matters, *Solutions émergentes aux défis du développement (volume 2)*.

# Partie I

## Phase conceptuelle



# Le choix des approches évaluatives

*Thomas Delahais*



## Introduction

Le mot « approche » fait partie de ces termes évaluatifs employés à tout bout de champ, mais dont la définition est difficile. Après tout, qu'y a-t-il de commun entre l'expérimentation par assignation aléatoire et l'évaluation féministe, mis à part qu'on les qualifie toutes les deux d'approches évaluatives ? Cependant, les approches sont des éléments utiles pour appréhender la façon de faire des évaluations, ancrées dans la théorie comme dans la pratique, et qui permettent de mieux se comprendre. Il est donc utile de s'en servir comme « blocs de construction » d'une évaluation (section 1). Par où commencer ? Nous proposons ici de considérer l'éventail des approches possibles en partant de leur point d'entrée : le choix des méthodes; l'interrogation des valeurs; l'utilité des travaux; l'exigence de justice sociale (section 2). Nous terminerons enfin par quelques mots visant à faciliter aux lectrices et aux lecteurs le choix d'une approche appropriée (section 3).

## Qu'est-ce qu'une approche évaluative ?

Une façon de définir les approches évaluatives est de les voir comme des modèles, comme un « ensemble de règles, de prescriptions et d'interventions, et de cadres qui guide [l'évaluateur(-trice)] et qui spécifie ce qu'est une bonne évaluation, ou une évaluation adéquate, et comment elle devrait être menée » (Alkin, 2013, p. 4). Ces approches peuvent être propres à l'évaluation ou bien être importées de la recherche, ou encore de pratiques voisines à l'évaluation.

## Un entre-deux, entre théorie et pratique

Une approche est une sorte d'entre-deux entre la théorie et la pratique. D'un côté, une évaluation s'inscrit, implicitement ou explicitement, dans une certaine vision du monde, de la nature des choses, de l'éthique, c'est-à-dire dans des paradigmes évaluatifs (Dagenais et Ridde, 2012, p. 14), qui se déclinent eux-mêmes en théories évaluatives. De l'autre, elle consiste aussi à manipuler des méthodes ou des outils des sciences sociales et à mener un certain nombre d'opérations très pratiques, de collecte d'informations par exemple. Une approche ne doit être confondue ni avec les uns ni avec les autres.

Ainsi, on peut décrire une évaluation en disant qu'elle se place dans un paradigme positiviste, c'est-à-dire qu'elle s'appuie sur l'idée qu'il existe une réalité qui est indépendante des perceptions de chacun de nous; et dans une théorie de la causalité contrefactuelle, au titre de laquelle on pourra affirmer qu'une cause C produit un effet E uniquement si, en l'absence de C, E est absent.

Inversement, on pourrait présenter cette évaluation par les actes concrets qu'elle suppose : des outils ou méthodes tels que l'enquête, voire à un niveau plus micro comme un échantillonnage, la définition d'indicateurs, des analyses statistiques, des choix d'interprétation, etc.

Mais même les plus acharnés des théoriciens et théoriciennes ne se posent pas de questions relatives à la nature de la réalité à chaque évaluation, et les plus terre-à-terre des évaluateurs ou évaluatrices peuvent comprendre qu'une suite d'outils ou d'actes ne suffit pas à rendre compte de ce qu'ils ou elles veulent faire. Une approche fait le lien entre ces deux extrêmes en fournissant de nombreux éléments permettant de mener à bien une évaluation (Tableau 1)<sup>1</sup>.

**Tableau 1 — Éléments constitutifs d'une approche**

Une approche fournit :	Afin de :
Un cadre paradigmatique, éthique et théorique.	Répondre à des questions relatives à la nature de la réalité et de la connaissance.
Une méthodologie (deMarrais et Lapan, 2004, p.5).	Décrire comment une évaluation devrait être menée en théorie pour qu'elle soit considérée comme étant de bonne qualité (chaque approche tendant à définir ses propres critères de qualité).
Des promesses et des situations d'usage préférentielles.	Choisir quand déployer la méthodologie, selon le contexte et les conséquences attendues de l'évaluation.
Un mode d'emploi (plus ou moins prescriptif), des éléments de posture et des concepts opératoires.	Mettre en œuvre cette méthodologie uniforme, indépendamment de la personne qui la déploie.
Une liste d'outils suggérés ou imposés, une façon de les articuler, parfois des indications particulières sur la manière de les utiliser.	Recueillir les informations nécessaires pour le bon fonctionnement de l'approche.

### Les approches comme communautés de pratique

Toutes les approches n'intègrent pas la totalité de ces éléments, ou plutôt, celles et ceux qui formulent une approche ne peuvent préciser l'ensemble de ces caractéristiques en une seule fois. Le ferment initial d'une approche, c'est le plus souvent une théorie dont on envisage les applications pratiques, une méthodologie dont on explicite le cadre théorique, ou une montée en généralisation de pratiques personnelles.

Dans tous les cas, pour qu'une approche soit complète et qu'elle puisse guider les choix évaluatifs, elle doit pouvoir compter sur une « communauté de pratique » (Wenger-Trayner et Wenger-Trayner, 2015, p. 1), qui va progressivement l'étoffer. Certes, certaines personnalités en sont venues à incarner des approches – que l'on pense à Michael Q. Patton avec l'évaluation axée sur l'utilisation (2008). En réalité, cependant, l'élaboration des approches est un processus dynamique. Au sein de communautés de pratique se tiennent des débats parfois vifs sur leurs différents éléments constitutifs. Les approches sont effectivement testées sur le terrain dans une logique d'essai-erreur ou d'hybridation,

1. Cette liste s'appuie notamment sur les cinq critères proposés par Robin Lin Miller (2010) pour examiner les théories évaluatives : la spécificité opérationnelle, le champ d'application, la faisabilité en pratique, l'impact discernable et la reproductibilité.

des leçons sont tirées de contextes nouveaux, de nouvelles solutions à des problèmes repérés, etc. Il ne suffit donc pas de décrire une méthodologie ou de fournir un cadre théorique pour qu'une approche prenne vie<sup>2</sup>.

Ainsi, lorsqu'une équipe d'évaluation souhaite s'engager dans une approche évaluative, par exemple une évaluation quasi expérimentale, elle peut s'appuyer à la fois sur des écrits qui dépeignent le déroulement d'une telle évaluation, mais aussi sur une série d'exemples auxquels se référer et de personnes-ressources auxquelles s'adresser. C'est le recours à ces exemples et à ces personnes qui permet à l'équipe d'utiliser concrètement une approche, en traitant de cas limites, de difficultés techniques ou éthiques, de biais repérés, etc.

Parfois, ces débats se font au sein d'une dénomination partagée. Mais très souvent, ils amènent à la constitution de familles d'approches. Ainsi, l'évaluation féministe regroupe sous son ombrelle plusieurs approches, qui ont pour point commun de s'inscrire dans une vision transformationnelle et qui se distinguent ainsi de l'approche attentive au genre (Podems, 2014). Inversement, on trouve autour de l'analyse de contribution de nombreuses propositions alternatives qui partagent des valeurs et des concepts similaires, et qui mentionnent toutes l'analyse de contribution dans leurs inspirations tout en s'en détachant<sup>3</sup>. Il peut être difficile de se retrouver dans ce foisonnement – d'où l'importance de communautés de pratiques et de personnes-ressources permettant d'y avoir accès – mais il participe à offrir la diversité de modalités qu'un évaluateur ou évaluatrice est en droit d'attendre d'une approche.

### **Les approches comme blocs de construction de l'évaluation**

Enfin, on peut définir les approches par leurs usages, et les voir comme des éléments préfabriqués, des blocs mis à la disposition des évaluateurs et évaluatrices pour construire une évaluation. Chaque évaluation nécessite l'élaboration d'une démarche évaluative ad hoc. Par ce terme, nous désignons l'ensemble des choix faits par l'équipe d'évaluation et visant à atteindre des buts évaluatifs. Ces choix sont de natures multiples : ils sont méthodologiques bien sûr, mais relèvent aussi de l'organisation, de la posture, de la relation aux parties prenantes, etc. Or, plus l'évaluation est complexe, plus ces choix sont marqués par l'incertitude : que faudrait-il faire, que peut-on attendre de telle ou telle démarche ?

Toutes les évaluations ne se positionnent pas explicitement dans des approches spécifiques, et ce n'est pas forcément un gage de qualité d'une évaluation qu'elle se réclame d'une approche. Il est d'ailleurs à noter que les pratiques varient entre les pays et les cultures d'évaluation à ce sujet : en Europe, il est sans doute moins fréquent de se réclamer d'une approche en particulier qu'en Amérique du Nord.

Cependant, la mention d'une ou de plusieurs approches évaluatives a plusieurs avantages :

- Les approches servent à ouvrir les choix possibles pour une évaluation, tout en offrant des cadres robustes et prévisibles. Avoir accès à un large répertoire d'approches est une occasion de sortir de ses habitudes ou de sa zone de confort;

---

2. Egon Guba et Yvonne Lincoln (1989), par exemple, ont proposé une théorie constructiviste de l'évaluation et ont décrit dans le détail ce que pourrait être une évaluation « de 4<sup>e</sup> génération ». Si la théorie s'est révélée très influente, en réalité elle n'a presque jamais été mise en pratique (Lay & Papadopoulos, 2007), et d'autres approches ou démarches s'inscrivant pour tout ou partie dans la théorie constructiviste ont prospéré.

3. On peut citer par exemple le Contribution tracing (Befani et Mayne, 2014), le Qualitative Impact Assessment Protocol (Copestake *et al.*, 2019), la méthode Impress (Barret *et al.*, 2017) ou l'Outcome Evaluation approach (Belcher, 2020).

- Les approches permettent de se placer dans une vision du monde, de ce qui est bien, de ce qui compte : expliciter ces valeurs permet de voir en quoi elles correspondent à ses valeurs personnelles ou à celles de l'organisation pour laquelle on travaille;
- Les approches facilitent la compréhension mutuelle entre les évaluateurs et évaluatrices, commanditaires, et autres parties prenantes de l'évaluation. Un choix commun d'approches est une assurance que certaines valeurs sont communes<sup>4</sup> – du moins si la connaissance des approches est partagée. Parler d'approches est rassurant (il existe d'autres exemples de ce type d'évaluation, qui ont été menés à bien) et, in fine, contribue à créer un climat de confiance utile à l'évaluation.

## L'éventail des approches

Note : dans cette partie, les références en français permettant d'en savoir plus sur les différentes approches présentées sont soulignées.

Dans tout l'éventail des possibles, nous avons privilégié les approches qui disposent aujourd'hui de communautés de pratiques actives, car ce sont les plus susceptibles de fournir aux équipes d'évaluation qui souhaiteraient les mettre en œuvre les ressources nécessaires pour le faire. Parmi ces approches, nous avons accordé une certaine importance à celles traitant des questions causales, car l'impact des interventions est au cœur de nombreux débats dans le monde de l'évaluation. Au-delà des choix méthodologiques<sup>5</sup>, l'impact est aussi une question d'approche, et c'est ce que nous illustrerons ici.

Toutes ces approches sont complexes et visent généralement à aborder différents aspects de la démarche évaluative. Il est donc difficile de les catégoriser sans tomber dans la superficialité. Nous avons choisi de recourir à quatre ingrédients de toute évaluation<sup>6</sup> : les méthodes, l'utilisation, les valeurs et l'attention à la justice sociale. Nous considérons que toute évaluation doit, dans des proportions qui varient selon les cas, s'inscrire dans ces quatre dimensions pour mériter son nom. Néanmoins, ces quatre branches correspondent à des « portes d'entrée » dans l'évaluation, qui nous semblent pertinentes pour naviguer dans les approches.

Dans le cadre de ce chapitre, la description des approches est nécessairement trop brève pour rendre compte de toutes leurs facettes. Nous avons voulu mettre en exergue leurs apports, leurs valeurs sous-jacentes et les contextes dans lesquels elles brillent particulièrement. Lorsque c'est possible, des ressources en français sont suggérées pour aller plus loin.

### Les approches orientées méthode

Les approches qui sont réunies dans cette partie accordent une importance particulière à la démarche d'investigation menée dans l'évaluation, ainsi qu'à la production de connaissances. C'est notamment ici que l'on retrouve les approches d'évaluation d'impact, qui visent à établir un lien de causalité entre

---

4. C'est d'autant plus important si, comme le soutient MQ Patton, « les gens recrutent probablement des évaluateurs parce qu'ils partagent des valeurs communes plutôt que pour les outils qu'ils proposent. » (cité dans Williams, 2018, p. 142).

5. Pour des listes de méthodes, voir Befani (2020) et Vaessen *et al.* (2020). Pour un aperçu des approches et méthodes en français, voir Quadrant Conseil (2017).

6. Nous reprenons ici la définition de l'évaluation en trois composantes proposées par Demarteau (2002, p. 457) en la complétant avec la préoccupation de justice sociale (Mertens et Wilson, 2012, p. 41).



une intervention et des changements observés. Nous distinguons dans un premier temps des approches qui se placent dans une logique expérimentale, où l'impact est la différence entre la situation avec et sans l'intervention, et qui sont d'abord préoccupées par la description de l'impact (Shadish *et al.*, 2002), et des approches dites basées sur la théorie, et en particulier les plus récentes, qui reconnaissent la complexité des contextes sociaux et qui cherchent en premier lieu à expliquer les conditions dans lesquelles les interventions ont des effets.

### Encadré 1 – Approches orientées méthode

- Les évaluations d'impacts
  - Le devis expérimental (l'expérimentation par assignation aléatoire)
  - Les devis quasi expérimentaux
  
- Les approches basées sur la théorie
  - L'analyse de contribution
  - L'évaluation réaliste

### Les évaluations d'impacts contrefactuels

#### Devis expérimental (ou l'expérimentation par assignation aléatoire)

L'expérimentation par assignation aléatoire (*randomized controlled trial*) est une approche « totale » qui dépasse l'évaluation pour se positionner comme outil de production de l'action publique par l'expérimentation. Par métonymie, cette approche est généralement appelée par le nom de sa méthode principale, l'essai contrôlé randomisé (ECR). L'ECR consiste à comparer, dans le cadre d'une expérimentation, les effets d'une intervention sur deux groupes, l'un des groupes y ayant accès et l'autre non, au regard d'un indicateur donné. Chaque groupe est tiré aléatoirement au sein de la population admissible, afin de prémunir l'expérimentation d'un biais de sélection. L'impact est alors la différence entre la valeur constatée pour le groupe avec l'intervention et le groupe sans intervention (White, 2014).

Au-delà de cet aspect, l'expérimentation par assignation aléatoire vise à créer les conditions pour juger « scientifiquement » du mérite intrinsèque d'une intervention hors influence de la mise en œuvre – le jugement étant positif si l'intervention marche dans ces conditions contrôlées. C'est le fait de fournir une preuve d'impact qui justifie, d'un point de vue éthique, le recours à un tirage aléatoire et les autres mesures mises en place – plutôt que d'aider ceux qui en auraient le plus besoin, par exemple.

Sous son apparente simplicité, la complexité de l'expérimentation par assignation aléatoire tient aux efforts extraordinaires nécessaires pour maintenir les conditions de la validité des résultats produits (Shadish *et al.*, 2012). Ces conditions limitent aussi son usage à des situations où un lien séquentiel direct entre une intervention et un effet peut être envisagé. L'expérimentation est conçue, le cas échéant, pour rendre cette situation observable.

### Exemple 1 – Expérimentation par assignation aléatoire

Une expérimentation est engagée auprès de jeunes diplômés qui peinent à trouver du travail. Pour en estimer les effets sur l'accès à l'emploi, un ECR est organisé. Il vise à comparer trois groupes, l'un « traité », l'autre de contrôle, et enfin un groupe de « super-contrôle », comprenant des jeunes issus de territoires non affectés par l'expérimentation. Cette modalité cherche à vérifier si l'accès au travail des jeunes inscrits se fait au détriment des autres jeunes diplômés (effet de substitution). L'analyse réalisée présente à première vue un impact mesurable sur l'emploi à 12 mois : à la fin de l'expérimentation, 39,5 % des jeunes concernés sont en emploi, contre 37 % des jeunes du groupe de contrôle. Cependant, ces taux sont inférieurs à ceux identifiés dans le groupe de super-contrôle. De plus, la reproduction des enquêtes à 16 mois et à 20 mois montre qu'il n'y a plus, à ce stade, de différences entre le groupe test et le groupe de contrôle (cité dans Behaghel *et al.*, 2013).

### Les devis quasi expérimentaux

Nous regroupons sous cette dénomination commune toutes les approches (*quasi-experimental designs*) qui visent à appliquer l'approche expérimentale, comparant une situation avec l'intervention, et une situation sans l'intervention, dans des contextes non expérimentaux. La comparabilité entre groupes ne pouvant être garantie à l'aide de l'assignation aléatoire, le groupe témoin doit être reconstitué, à partir de bases de données existantes ou produites pour l'occasion (par enquête, par exemple).

Des concepts et des méthodes statistiques sont utilisés pour garantir la comparabilité des groupes, en s'assurant de la proximité de la composition entre les deux groupes et/ou en prenant en compte leurs différences dans l'estimation de l'impact. Dans le premier cas, on utilise des méthodes d'appariement : pour chaque individu du groupe test, on trouve un « jumeau » (c'est-à-dire un individu aux caractéristiques observables similaires), en utilisant des techniques plus ou moins sophistiquées pour ce faire. Dans le second, on mesure les différences pour les retirer de l'opération : la méthode de double différence consiste ainsi à comparer deux groupes, l'un soumis à l'intervention, l'autre non, et ce, avant, à l'entrée, et à la sortie du dispositif, au regard d'un indicateur d'impact. Les niveaux mesurés sont différents, mais si les groupes sont suffisamment similaires, leurs évolutions le sont également. L'impact est estimé en défalquant de la différence entre les deux groupes à la sortie du dispositif, celle observée à l'entrée (White et Sabarwal, 2014).

### Exemple 2 – Devis quasi expérimental

Un programme vise à lutter contre la malnutrition infantile au Bangladesh à travers des ateliers ciblant les mères de famille. Pour en vérifier les effets, des données relatives à l'âge, à la taille et au poids des enfants inclus dans le dispositif sont comparées à celles réunies dans le cadre d'une enquête nutritionnelle alimentaire. Un appariement par score de propension, « qui désigne la probabilité de participation à l'intervention au vu des caractéristiques observables de l'individu » (White et Sabarwal, 2014) permet de construire le groupe de contrôle. L'impact de l'intervention est ensuite calculé en comparant les chiffres obtenus en moyenne pour chaque groupe. L'évaluation montre que, contrairement à ce que semblaient montrer les données de suivi, les enfants passés par le programme ne se distinguent pas significativement des autres enfants bangladais à caractéristiques comparables (White, 2009).

## Les approches basées sur la théorie

Nous regroupons sous cette ombrelle commune (*theory-based evaluation*) une série d'approches partageant des caractéristiques communes, et relevant d'une volonté d'expliquer la façon dont les interventions évaluées ont des effets (Brousselle et Buregeya, 2018). Elles sont appelées ainsi parce qu'elles organisent toutes ce processus explicatif autour de la constitution puis de la vérification d'une théorie du changement (TdC) (Funnell et Rogers, 2011)<sup>7</sup>. Cette théorie « avec un petit t » (Chen et Rossi, 1987) est élaborée au début de l'évaluation à partir d'hypothèses implicites ou explicites portant sur l'intervention et sa capacité à avoir des conséquences, le pourquoi et le comment, et dans quels cas. Des questions d'évaluation sont ensuite posées pour interroger tout ou partie des hypothèses réunies dans la TdC, et des outils de collecte et d'analyse sont choisis et articulés pour ce faire. L'évaluation se termine par une mise à jour de la théorie, qui est mise à contribution pour améliorer la faculté de l'intervention à avoir des effets désirables.

Un aspect fondamental de ces approches est de considérer que les interventions ne « marchent pas » toutes choses égales par ailleurs : elles reconnaissent au contraire une diversité de cas, et considèrent le « contexte » comme un élément affectant les changements observés. C'est par ce biais qu'elles abordent les interventions comme des objets complexes dans des environnements complexes (Robert et Ridde, 2013).

Différentes approches distinctes prospèrent aujourd'hui au sein de cette famille. Nous en explorons deux ci-dessous.

### L'analyse de contribution

L'analyse de contribution (*contribution analysis*) se distingue dès ses premiers pas par une démarche humble et pragmatique : son objectif n'est pas de prouver les effets des interventions, mais bien de réduire l'incertitude quant à leur contribution, de façon à éclairer la décision. C'est ainsi qu'elle est progressivement devenue une approche particulièrement adaptée aux situations complexes, lorsqu'une contribution est plausible, mais n'est certainement ni séquentielle ni linéaire (Mayne, 2012), ou lorsqu'il est probable, dès le départ, que l'intervention ne soit pas le principal vecteur du ou des changements recherchés (Delahais et Toulemonde, 2017).

Une originalité de l'analyse de contribution est que sa recherche de liens causaux ne se limite pas à l'intervention. Au contraire, elle vise spécifiquement à repérer des « paquets de causes », associant l'intervention et d'autres facteurs, et qui peuvent être des conditions nécessaires ou suffisantes aux changements escomptés.

La démarche d'analyse de contribution peut être assimilée à une sorte d'enquête : à partir de la théorie du changement initiale, des hypothèses de contribution sont bâties, et leur plausibilité testée de diverses façons. Ces hypothèses portent à la fois sur l'intervention et sur des explications alternatives. Elles sont ensuite testées, à charge et à décharge, lors de la collecte proprement dite. Dans ce processus, mesure des changements et estimation des contributions sont strictement séparées, ce qui demande à la fois un travail de documentation des changements et d'élaboration de tests empiriques relatifs à

7. Il existe de très nombreux termes qui peuvent être utilisés pour parler de chaînes causales représentant les interventions et leurs conséquences de façon séquentielle, nous utilisons ici ce terme de façon générique. Pour une vue générale, voir Funnell et Rogers (2011).

la contribution. L'analyse de contribution fournit un effort particulièrement important d'identification des impacts en contexte; en contrepartie, elle ne prétend pas à établir des modèles généraux, mais plutôt les ingrédients et les circonstances pour leur utilisation (Delahais *et al.*, 2020).

### Exemple 3 – Analyse de contribution

Un centre de recherche souhaite évaluer sa contribution à la gestion durable des forêts dans le bassin du Congo sur les 20 dernières années. Il est évident qu'il n'en est pas le principal vecteur, mais les parties prenantes font valoir des contributions plausibles de la recherche qui peuvent être approfondies. Un système représentant les principaux acteurs (États, organisations internationales, acteurs de la recherche, ONG, entreprises, habitant(e)s...) ayant contribué à cette gestion durable est élaboré à partir de la littérature, et permet de documenter les principaux changements intervenus sur la période. Des hypothèses relatives à la contribution du centre de recherche aux différentes étapes de ce système sont émises, puis testées systématiquement au travers d'études de cas nationales et thématiques : in fine, l'évaluation fait apparaître sept contributions majeures de la recherche à la gestion durable des forêts, à plusieurs points du système, articulées avec d'autres facteurs expliquant ces changements (Delahais et Toulemonde, 2017). Par exemple, le centre de recherche fournit les connaissances nécessaires pour pouvoir développer une certification durable adaptée aux forêts de la région, elle-même nécessaire pour que les entreprises puissent mieux valoriser le bois produit de cette façon.

### L'évaluation réaliste

L'approche réaliste (*realist evaluation*) tire son nom du paradigme réaliste critique dont elle est l'adaptation à l'évaluation (Pawson et Tilley, 1997). Elle se distingue ainsi par sa conceptualisation des effets, son objet et sa volonté de généralisation.

Dans l'approche réaliste, en effet, les interventions ne « marchent pas » d'elles-mêmes : elles fournissent à des acteurs une opportunité d'agir d'une certaine façon, en raison de leurs caractéristiques propres, des interactions qu'ils ont entre eux et avec les acteurs qui portent l'intervention, et selon le moment ou les circonstances. Les raisonnements multiples qui les amènent à se saisir ou non de l'intervention sont appelés « mécanismes. » De ceux-ci dépendent in fine les effets de l'intervention, et l'évaluation s'attache à les repérer et à comprendre les contextes dans lesquels ils sont activés ou non.

Là où l'analyse de contribution cherche à établir des systèmes, l'évaluation réaliste cherche plutôt à épuiser les explications relatives à un moment donné de la théorie du changement. Elle est notamment utilisée pour comprendre la façon dont sont mises en œuvre les interventions. Ces explications sont présentées sous la forme de configurations CME : dans tel Contexte, une intervention déclenche tel Mécanisme qui génère tel Effet (en anglais CMO, pour *Context-Mechanism-Outcome*). L'analyse méthodique de ces configurations doit aboutir à des théories de moyenne portée, qui permettent d'expliquer les régularités observées dans les configurations CME (Robert et Ridde, 2013), et ainsi, in fine, d'améliorer les interventions.

### Exemple 4 – Évaluation réaliste

Le programme *Dance with Parkinson's* consiste en des cours de danse pour les personnes atteintes de cette maladie dégénérative. Cette intervention est un succès : les participants et participantes reviennent régulièrement. Ses promoteurs y ont apporté des améliorations continues, mais rencontrent des difficultés à maintenir leurs financements, et cherchent alors à prouver de façon systématique les effets obtenus. L'approche évaluative mise en place est participative et réunit les acteurs du programme pour construire une TdC, dans laquelle sont précisés les changements attendus sur l'état des participants et participantes, et les chemins par lesquels l'intervention peut exercer une influence sur cet état. Des observations et des entretiens avec les différentes personnes impliquées (enseignant(e)s, participant(e)s, leur famille) permettent de tester cette théorie. Plusieurs mécanismes sont identifiés, liés notamment à un sentiment de liberté retrouvée, à la confiance en soi, au plaisir également (les classes sont des moments amusants), qui génèrent des effets à court terme (meilleure humeur, moins de chutes...) et à long terme (meilleure qualité de vie pour les participants et participantes et pour ceux qui leur prodiguent des soins) (Gibson et Robichaud, 2020).

### Les approches concernées par les valeurs

Nous proposons ici une série d'approches dans lesquelles ce sont les valeurs des parties prenantes qui sont au cœur de l'évaluation. Quelles que soient leurs différences par ailleurs, ces approches prévoient d'identifier les parties prenantes et leurs valeurs, et organisent l'évaluation de façon à examiner l'intervention au regard de celles-ci. Aussi ces approches comportent-elles toujours une dimension participative ainsi que des outils de collecte et d'analyse permettant de rendre compte finement des contextes, des attitudes et des comportements des acteurs.

### Encadré 2 – Approches concernées par les valeurs

- La famille des approches participatives
- L'évaluation répondante
- L'évaluation démocratique délibérative

### La famille des approches participatives

Les approches présentées dans les pages suivantes, qu'elles soient concernées par les valeurs, par l'utilisation ou par la justice sociale, accordent une place importante aux parties prenantes. Elles peuvent généralement être menées de façon participative (mais pas toujours).

Une évaluation participative donne à un groupe de parties prenantes un rôle décisionnel dans la conduite de l'évaluation, ce contrôle pouvant être total ou partagé; ces parties prenantes représentent une diversité de perspectives sur l'intervention, et à tout le moins d'autres points de vue que ceux des décideurs ou des financeurs; enfin, les parties prenantes doivent contribuer à la production de l'évaluation et de ses cadres, au minimum en influant sur la façon dont l'intervention va être jugée et sur ses usages. Le processus évaluatif organise la délibération entre les parties prenantes, permettant de s'informer, mais aussi d'évoluer dans ses préférences et d'aboutir à un jugement plus fondé (Delahais, 2021).

On peut considérer deux visées principales des évaluations participatives (Cousins et Whitmore, 1998). D'un côté, produire de meilleures évaluations, plus pertinentes, plus crédibles et plus utiles. C'est ainsi que les différentes approches concernées par l'utilité de l'évaluation voient en priorité la participation. De l'autre, donner à ceux qui n'ont pas la parole habituellement les moyens de se faire entendre et de changer les choses. C'est la perspective des approches orientées vers la justice sociale. Les deux approches présentées ci-dessous se placent entre ces deux visées, avec un souci d'être utile, mais sans méconnaître les enjeux de pouvoir et de domination qui affectent la société.

### L'évaluation répondante

La prémisse essentielle de l'évaluation répondante (*responsive evaluation*) est de considérer que les parties prenantes ont des valeurs différentes; que ces valeurs jouent un rôle dans la façon dont les uns et les autres jugent une intervention; que l'évaluateur ou l'évaluatrice ne devrait pas imposer de valeurs particulières dans le jugement apporté. Cette ascendance constructiviste étant établie, tout le rôle de l'équipe d'évaluation est de construire le processus permettant de déterminer les valeurs qui comptent, de faire émerger les enjeux qui intéressent les différents acteurs impliqués pour pouvoir y répondre et de favoriser, in fine, une pluralité de jugements. En ce sens, l'évaluation répondante cherche à aider ces derniers à se faire leur avis sur l'intervention, de façon à faire les choix qui leur semblent adéquats.

Un des défis essentiels de l'évaluation répondante est de se donner les marges de manœuvre suffisantes pour identifier les sujets qui feront effectivement l'objet de l'évaluation. Au début de l'évaluation, les parties prenantes ont des questions relativement basiques, du type « est-ce que les objectifs sont atteints? ». Or, l'évaluation répondante ne donne pas une place prépondérante aux objectifs (ou à la logique d'intervention) dans la définition de ce qui compte. Il faut du temps et de la flexibilité, des interactions avec l'équipe d'évaluation, des premiers éléments de collecte pour que des questionnements plus essentiels émergent et guident la suite de l'évaluation. Les études de cas, qui explorent en profondeur une situation, son contexte, les effets attendus et inattendus d'une intervention, les différents points de vue à son sujet, sont l'outil privilégié de l'approche répondante, non seulement pour la richesse d'informations qu'elles apportent, mais aussi parce qu'elles permettent aux parties prenantes d'appuyer leur jugement.

Dans cette approche, les interactions entre les parties prenantes et l'équipe d'évaluation sont nombreuses, mais les rôles sont clairement répartis : c'est l'équipe d'évaluation qui, à partir des interrogations des acteurs, mène ses investigations et en rapporte ensuite les résultats. Cependant, les acteurs de l'intervention sont considérés comme des participants actifs à l'enquête, plutôt que des sujets : on leur demande comment ils perçoivent les changements, plutôt que de leur faire passer des tests pour les mesurer (Stake, 2003).

### Exemple 5 – Évaluation répondante

Un programme aide les anciens patients et patientes d'un établissement psychiatrique à retrouver un emploi. Ce programme se construit chemin faisant, et l'évaluation vise à permettre à ses acteurs, et en particulier aux personnels, à réfléchir à leurs pratiques et à les améliorer. Très rapidement, il apparaît à l'évaluatrice que les pratiques des thérapeutes sont probablement un obstacle à la réinsertion des patients et patientes. Dans leur vision des choses, ces derniers ont besoin de l'expertise, des soins et de la protection des personnels; ils ne savent pas ce qui est bon pour eux. Pour faire évoluer leur pratique, il faut faire changer ce récit. L'évaluatrice va alors collecter d'autres histoires, qui ne cadrent pas avec ce récit dominant, auprès de thérapeutes, de patients et patientes, et de spécialistes, et les rapporte aux parties prenantes. Elle montre comment les patients et patientes vivent ces pratiques et comment elles les maintiennent dans la dépendance. Ce processus amène certains personnels à se rendre compte que leur pratique pose problème, et à engager un processus de transformation. Le rapport ne comprend ni conclusions (à chacun d'en tirer) ni recommandations (Abma, 1998).

### L'évaluation démocratique délibérative

Cette approche (*democratic deliberative evaluation*) a l'ambition de contribuer à la bonne marche de la démocratie, en faisant la démonstration de ce qu'est un processus démocratique, mais aussi en permettant que les décisions prises dans le cadre des institutions démocratiques le soient à partir de constats robustes et de conclusions étayées. Pour cela, trois principes majeurs sont mis de l'avant : l'inclusion des parties prenantes; le dialogue; la délibération.

L'évaluation démocratique partage l'idée que les interventions sont l'objet de nombreuses perspectives, mais elles ne sont pas toutes égales : pour être équitable, l'équipe d'évaluation doit rechercher activement la participation de certaines parties prenantes et doit éviter que les plus puissants n'imposent leurs vues. Ces parties prenantes doivent pouvoir avoir l'occasion de dialoguer sur les enjeux identifiés et les analyses apportées par l'évaluation. Ce dialogue permet à l'équipe d'évaluation de comprendre pleinement les positions et les points de vue des groupes en présence, et à ceux-ci de clarifier leurs propres intérêts vis-à-vis de l'intervention. C'est grâce à ce processus que sont définis les principaux sujets d'investigation. Le processus de délibération vise enfin à mettre en perspective ces différents points de vue avec les informations collectées lors de l'évaluation : il ne s'agit pas de les prendre pour argent comptant, et l'objectif de l'évaluation est bien d'aboutir à un jugement de la valeur de l'intervention qui peut être acceptable pour le plus grand nombre. C'est là où l'équipe d'évaluation doit prendre ses responsabilités et proposer, in fine, les conclusions qui lui semblent les plus adaptées, les débattre avec les parties prenantes (désormais averties de la diversité des perspectives et des éléments les soutenant ou pas) et les consolider (House et Howe, 2013).



### Exemple 6 – Évaluation démocratique délibérative

Un programme fournit une éducation bilingue à des élèves qui ne maîtrisent pas suffisamment la langue officielle d'enseignement. Le contexte, dans cette ville américaine, est difficile, avec une animosité croissante entre communautés, qui se cristallise autour de ce programme. L'évaluateur commence par rencontrer les parties prenantes, y compris les principaux soutiens et adversaires du programme. Des mesures sont ensuite prises pour s'assurer que tous les groupes en présence acceptent la démarche évaluative et ses résultats : ainsi, un comité composé des différentes parties est réuni; les indicateurs de performance sont discutés jusqu'à ce qu'ils soient unanimement considérés comme des indicateurs de succès; la collecte est menée par deux anciens directeurs d'établissement, respectés de tous, etc. L'évaluateur accorde une attention particulière aux affirmations des parties. Par exemple, certaines familles se plaignent que les élèves sont envoyés trop tôt en classe d'anglais : le niveau des élèves lors du passage d'une classe à l'autre fait donc l'objet d'investigations spécifiques. Se basant sur cinq ans d'évaluation, l'évaluateur met de l'avant que le programme fonctionne désormais correctement, et que les parties prenantes peuvent échanger entre elles, et considère que le processus mis en place, transparent, équitable, juste, y est pour quelque chose (Mertens et Wilson, 2012)<sup>8</sup>.

### Les approches concernées par l'utilisation

Nous regroupons ici des approches foncièrement concernées par l'utilisation de l'évaluation, et qui sont organisées en conséquence. Michael Patton a une importance singulière, tant pour son investissement de longue date sur ce sujet, que pour sa capacité à formaliser ses pratiques évaluatives sous la forme de modèles complets. Nous commencerons donc par deux d'entre elles, et terminerons par une troisième approche, la récolte des incidences, qui se place dans la filiation de Patton tout en s'en démarquant par une méthodologie originale.

### Encadré 3 – Approches concernées par l'utilisation

- L'évaluation axée sur l'utilisation (ÉAU)
- L'évaluation évolutive
- La cartographie et la récolte des incidences

8. Cette évaluation a été menée par E. House lui-même. Elle est résumée dans Mertens et Wilson (2012, p. 176-178).



## L'évaluation axée sur l'utilisation (ÉAU)

Lorsqu'elle apparaît sur la scène évaluative, cette approche (*utilization-focused evaluation*) suscite la controverse par une proposition essentielle : « les évaluations devraient être jugées en fonction de leur utilité » (Patton, 2008; Patton et Labossière, 2012), et non en fonction de la méthodologie déployée, par exemple. L'utilisation n'est pas quelque chose que l'on attend vaguement à la fin de l'évaluation, mais l'objet même de toute la démarche. L'ÉAU n'est ainsi pas prescriptive en matière d'outils de collecte ou d'analyse, l'important est que ceux-ci permettent de répondre aux besoins d'information affirmés par les usagers.

Un des éléments essentiels de l'ÉAU est qu'elle est entièrement organisée autour des personnes qui vont être impliquées dans le processus évaluatif. Elle démarre donc par l'identification des principaux utilisateurs attendus d'une évaluation, qu'il s'agisse des commanditaires ou d'autres acteurs, et la détermination de leur niveau de participation à la démarche évaluative. Les parties prenantes s'accordent ensuite sur les utilisations prévues de l'évaluation (quels besoins d'évaluation, pour quoi faire, à quel moment) et précisent le questionnement évaluatif en conséquence. Les utilisateurs sont amenés à s'impliquer à la fois dans le choix des outils mis en œuvre pour répondre aux questions, puis dans l'interprétation des résultats obtenus, toutes ces étapes visant à renforcer leur engagement et in fine l'utilisation de l'évaluation. Ainsi, le déroulement de l'évaluation est, tout autant que les résultats, vecteur de meilleure utilisation.

L'ÉAU fait de l'évaluation un processus dynamique, susceptible d'évoluer en cours de route pour préserver, renforcer ou élargir le potentiel d'utilisation. Elle donne un rôle crucial d'animation et de négociation à l'évaluateur ou à l'évaluatrice, qui doit en plus intégrer des enjeux éthiques plus importants qu'à l'accoutumée, relatifs aux usages qui seront faits de l'évaluation.

### Exemple 7 – Évaluation axée sur l'utilisation

Deux expérimentations visent à accueillir des jeunes en grande précarité en résidence (sous forme d'internat ou de séjour pendant les vacances scolaires) pendant leur scolarité. Dans un premier temps, l'équipe d'évaluation échange sans démarche évaluative prédéfinie avec des utilisateurs potentiels. La construction du référentiel est ensuite engagée de façon itérative, elle permet d'identifier une problématique et des questions d'évaluation intéressant toutes les parties ainsi que les principaux éléments de méthode pour y répondre. Les échanges font apparaître le besoin de mettre en place un système de suivi, pour aider les opérateurs à ajuster et à structurer l'action au fur et à mesure : le choix des indicateurs fait l'objet de négociations avec les utilisateurs potentiels, dans l'optique de trouver un bon équilibre entre ce qui serait le plus utile et ce qui est possible de faire pour suivre sans nuire au travail des équipes. Le travail sur les écarts entre le cadre de vie des jeunes et celui proposé dans les structures d'accueil (culture, discipline, modes d'interaction, confort) s'avère assez rapidement un enjeu clé de la réussite du projet sur le long terme, et les porteurs de projet demandent un cadre leur permettant d'y réfléchir : l'équipe d'évaluation identifie et propose alors des concepts théoriques qui serviront à structurer l'évaluation ainsi qu'à outiller les opérateurs. Les échanges lors de l'évaluation, y compris sur la collecte effectuée et l'interprétation des données, favorisent l'appropriation de l'évaluation par tous et toutes. In fine, des résultats distincts sont fournis aux différentes catégories d'utilisateurs (Devaux-Spatarakis *et al.*, 2021).

## L'évaluation évolutive

Contrairement à l'ÉAU, qui s'apparente presque à une posture pouvant s'appliquer à tout contexte évaluatif, l'évaluation évolutive (*developmental evaluation*) est conçue pour maximiser l'utilisation de l'évaluation dans le cadre très précis de l'innovation sociale, ou plus largement de changements radicaux qui induisent une très forte incertitude quant à leur mise en œuvre et à leurs impacts. Ces innovations fonctionnent par itérations successives, qui peuvent être très fréquentes. Or, une évaluation classique prend du temps pour poser ses critères et collecter des informations pour renseigner ces derniers; elle risque d'être systématiquement trop tardive pour être utile – ou pire, inadéquate – par exemple parce qu'elle travaille sur des hypothèses qui sont déjà obsolètes.

Avec l'évaluation évolutive, l'évaluateur ou l'évaluatrice intègre les équipes chargées de l'intervention évaluée, et s'attache à apporter des rétroactions rapides et pertinentes pour permettre aux innovateurs d'apprendre en temps réel et de mener à bien leurs ambitions. La posture évaluative en est considérablement modifiée : l'évaluatrice ou l'évaluateur est aussi intéressé que les autres à la réussite de l'intervention et agit en conséquence. Elle ou il met à profit sa pensée critique et sa posture évaluative, notamment pour cadrer les enjeux émergents, identifier des changements en cours, construire des typologies, poser des hypothèses et mettre en place des moyens pour les vérifier. L'évaluation doit aussi être en position de passer d'un rôle à un autre dans sa relation avec les parties prenantes : membre de l'équipe, ami critique, accompagnateur de changements, communicant, le tout demandant une forte flexibilité et des qualités dépassant largement celles qui sont généralement associées à l'évaluation (Dozois *et al.*, 2011; Turcotte *et al.*, 2021).

### Exemple 8 – Évaluation évolutive

De plus en plus, les multinationales du numérique délocalisent certaines de leurs activités dans des pays du sud. Un appel de projets d'une Fondation états-unienne est engagé au Kenya, en Afrique du Sud, au Ghana et en Inde pour vérifier si des actions de formation des jeunes leur permettraient de mieux bénéficier de ces occasions d'emploi. L'équipe d'évaluation, engagée en amont, s'attache à partager ses objectifs de créer de la connaissance partagée et de contribuer à l'amélioration collective, de façon à éviter la défiance qu'une démarche plus classique aurait pu faire naître. Elle multiplie les échanges, et répond aux besoins des porteurs et porteuses quand ils se manifestent. Rapidement, l'équipe d'évaluation est considérée comme faisant partie de l'équipe projet, ce qui facilite la transmission d'informations; inversement, les porteurs et porteuses intègrent les questions d'évaluation dans leurs processus et s'attachent à contribuer à l'acquisition de connaissances à travers leurs actions (Harji et Jackson, 2016).

## Les approches de cartographie et récolte des incidences

Ces deux approches, distinctes, mais partageant de nombreuses caractéristiques, représentent un continuum allant du suivi à l'évaluation, jusqu'à la mise en œuvre des interventions. Leur originalité est qu'elles naissent d'une définition spécifique des incidences (*outcomes*) : « des changements dans le comportement, les relations, les actions, les activités, les politiques, ou les pratiques d'un individu, d'un groupe, d'une communauté, d'une organisation ou d'une institution » (Earl *et al.*, 2001). Cette définition met les acteurs au centre des réflexions (et des processus, souvent participatifs); tient à

distance les objectifs ou autres résultats prédéterminés; et s'ouvre au contraire à l'ensemble des conséquences, attendues ou non, désirables ou non, des interventions. Pour cela, elle est particulièrement intéressante dans des contextes où ce sont les conséquences dans leur globalité qui importent (par exemple dans l'action sociale ou culturelle), ou encore lorsque l'évaluation porte sur des grappes de projets aux effets très différents, qui ne pourraient pas être captés par une théorie du changement.

La cartographie des incidences (*outcome mapping*) vise, en amont, à identifier les incidences et à construire des instruments de planification et de mesure autour d'elles. C'est un processus itératif, qui peut tout autant être mobilisé pour le suivi qu'à la mise en œuvre de l'intervention. La récolte des incidences (*outcome harvesting*) part, elle, en aval, de changements concrètement observés et cherche, à partir de là, à repérer les contributions de l'intervention évaluée. Elle peut se faire en temps réel, et contribuer ainsi à faire évoluer les interventions, ou bien servir à réaliser une évaluation rétrospective.

L'utilisation est au cœur de l'ensemble de ces démarches. La 1<sup>re</sup> étape de la récolte des incidences consiste ainsi à identifier avec les utilisateurs des questions qui répondent à leurs besoins, et la dernière à organiser une discussion entre « récolteurs » d'information et utilisateurs. Les processus sont conçus pour être simples d'accès. L'accent sur les personnes permet d'impliquer plus facilement toutes les parties prenantes, en particulier celles qui ne se retrouveraient pas dans des objectifs jugés artificiels (Wilson-Grau et Britt, 2012).

### Exemple 9 – Récolte des incidences

Une organisation à but non lucratif accompagne des communautés locales dans la mise en place d'actions en faveur du développement des tout-petits et du bien-être de leurs familles. Elle met en place une récolte des incidences portant sur ses soutiens apportés aux groupements de partenaires locaux pour juger rétrospectivement de son intervention. Une partie de la démarche est réalisée par entretiens individuels auprès de 20 groupements, permettant de recueillir 223 incidences, dont 24 négatives. Elles sont ensuite regroupées et analysées selon plusieurs variables (axes stratégiques, temporalité, durabilité, thématiques d'interventions, etc.) pour rendre compte des apports du programme. Parallèlement, des ateliers participatifs de récolte sont organisés dans les regroupements qui le souhaitent et ils visent à consolider leurs apprentissages dans une perspective de pérennisation des acquis après dix ans de soutien. Les participants et participantes présentent des changements qu'ils ont observés, donnent un exemple concret, exposent les conditions de succès nécessaires pour produire les changements et le faire perdurer, et réfléchissent au lien avec les contributions de l'organisation qui les a soutenus<sup>9</sup>.

### Les approches concernées par la justice sociale

Les approches présentées ici mettent l'évaluation au service d'une société plus juste et plus équitable. Nous parlerons dans un premier temps des approches transformatives, avant de terminer par une approche très différente, l'évaluation émancipatrice.

9. Cet exemple est tiré d'une expérience réelle, menée au Québec entre 2017 et 2019 par l'organisation Avenir d'Enfants. Pour en savoir plus : <https://agirtot.org/thematiques/recolte-des-effets-1-de-2/>

#### Encadré 4 – Approches concernées par la justice sociale

- Les approches transformatives
- L'évaluation féministe
- L'évaluation attentive aux différences culturelles et l'évaluation autochtone
- L'évaluation émancipatrice

#### Les approches transformatives

Ces approches sont inspirées de courants critiques tels que la recherche-action participative, les luttes anti-discriminations ou le féminisme. Ces approches ont des proximités avec les approches centrées sur les valeurs, mais s'en distinguent en cela qu'elles nomment clairement les maux qu'elles combattent (le racisme, le sexisme, et toute forme de discrimination et d'oppression) et cherchent à donner aux groupes n'ayant pas habituellement voix au chapitre les moyens de se faire entendre en prenant le pouvoir sur les processus évaluatifs.

L'évaluation est donc une activité politique. Cela veut dire qu'elle est explicitement mise au service d'une lutte pour la justice sociale (et doit donc être utile si elle veut faire évoluer le statu quo). Mais cela signifie aussi qu'elle est elle-même inscrite dans des relations de pouvoir et des inégalités structurelles. L'évaluatrice ou l'évaluateur doit donc être conscient de la place occupée, de son pouvoir, de la nature de ses savoirs (légitimés par le modèle dominant) et de ses lunettes culturelles, pour pouvoir les mettre à distance; chercher constamment à organiser les conditions de la meilleure prise en compte de l'expertise des groupes vivant les inégalités; avoir un sens intense de l'éthique; tout cela en gardant en permanence à l'esprit le souci de la transformation sociale, et les chemins par lesquels l'évaluation peut y contribuer.

Les approches transformatives ne sont pas prescriptives en termes d'étapes, mais Donna Mertens propose le déroulement suivant : une phase initiale qui doit permettre de construire une évaluation avec un vrai potentiel de transformation, et dans laquelle il faut donc connaître les publics, créer avec eux des relations de confiance et trouver les façons de faire bon usage de leur expertise; une phase de collecte d'informations qui peut utiliser une large variété d'outils, selon les besoins; une phase d'interprétation collective des données collectées, avec un objectif de construction des usages de l'évaluation (Mertens, 2017).

L'évaluation féministe et l'évaluation attentive aux différences culturelles, présentées ci-dessous, s'inscrivent dans la famille des approches transformatives.

#### L'évaluation féministe

Cette approche (*feminist evaluation*) prend comme point de départ les rapports de genre et avance l'idée que ceux-ci génèrent des injustices sociales qui sont systématiques et structurelles. En « chaussant des lunettes de genre », en analysant les rapports de genre et leurs conséquences, comment ils affectent l'accès aux interventions évaluées, leur mise en œuvre ou leurs effets, l'équipe d'évaluation effectue un pas de côté qui lui permet de voir l'intervention différemment. À noter que malgré son choix

d'appellation, l'évaluation féministe est aussi sensible aux autres sources d'injustice et cherche à comprendre comment ces facteurs s'additionnent et se renforcent mutuellement (intersectionnalité) (Podems, 2018).

### Exemple 10 – Évaluation féministe

Un programme s'adressant à des personnes toxicomanes et présentant des troubles psychiatriques leur propose un module optionnel visant à leur faire prendre conscience des traumatismes qu'elles ont vécus. Les évaluatrices ont repéré que les femmes étaient très peu nombreuses à s'inscrire à ce module et ont donc intégré un cadre féministe à leur évaluation. Cela leur a permis de mettre en regard le discours des thérapeutes, qui expliquaient initialement le faible taux de recours par le caractère facultatif du module et ses exigences en matière de temps, de réflexion et de travail personnel, avec l'expérience des femmes. Celles-ci, en effet, s'inscrivent moins, car elles déclarent moins que les hommes les traumatismes qu'elles ont vécus, et ce pour plusieurs raisons : 1) les femmes ne font pas toujours le lien entre leur santé mentale et leur expérience traumatique; 2) pour les femmes, la culture rurale rend plus difficile la reconnaissance qu'elles puissent avoir besoin d'aide; 3) les thérapeutes ont tendance à ne considérer comme traumatiques que les abus physiques et sexuels. C'est la réflexivité sur la manière dont les valeurs personnelles des évaluateurs et évaluatrices « s'alignent, s'entremêlent et/ou entrent en conflit avec les valeurs des principales parties prenantes lorsqu'elles définissent les problèmes, identifient les populations cibles, conçoivent les programmes, mesurent les résultats et rendent compte des conclusions » qui permet de produire des connaissances capables de générer du changement (Sielbeck-Mathes et Selove, 2014).

L'évaluation féministe, approche transformationnelle, ne doit pas être confondue avec les approches attentives au genre (*Gender-Responsive Evaluation*), qui visent plutôt à évaluer les interventions au regard d'un objectif d'égalité des sexes et d'autonomisation des femmes, sans remettre en question leur dimension structurelle (ONU Femmes, 2015).

### L'évaluation attentive aux différences culturelles et l'évaluation autochtone

Cette approche transformationnelle (*culturally responsive evaluation*) part du principe que « des valeurs et des croyances liées à la culture sont au cœur de tout effort évaluatif » (Hood *et al.*, 2015, p. 283), la culture étant définie comme « une accumulation de comportements, valeurs, coutumes ou croyances, apprises et partagées, qui sont communs à un groupe ou une société en particulier » (idem). L'évaluation attentive aux différences culturelles porte une attention particulière aux groupes culturels marginalisés ou discriminés en raison de leur origine, de leur apparence physique, de leur capacité à s'exprimer dans la langue officielle ou usuelle d'un pays, ou de leur situation économique ou ethnique, parmi tant d'autres.

Les postulats de cette approche évaluative ont plusieurs implications majeures. La première, c'est la nécessité de repérer les groupes culturels qui sont concernés ou touchés par les interventions évaluées – sachant que chaque individu a des identités multiples, il y a donc un travail important à faire pour comprendre lesquels forment une typologie pertinente dans un contexte donné. La deuxième, c'est la prise en compte des valeurs et des croyances liées à la culture de ces groupes dans l'ensemble de la démarche évaluative : questionnement évaluatif autour d'enjeux qui préoccupent et affectent ces

groupes; conception et mise en œuvre d'outils permettant de faire ressortir les modes de connaissance propres à ces groupes, et favorisant leur investissement dans l'évaluation; interprétation des résultats à l'aune des modes de connaissance de ces groupes. La troisième et dernière implication est la nécessité pour l'évaluateur ou l'évaluatrice d'avoir ou de développer des compétences culturelles qui sont appropriées pour le groupe culturel concerné. Outre des savoirs, savoir-faire et savoir-être que tout évaluateur ou évaluatrice peut acquérir, il s'agit aussi, à chaque nouvelle évaluation, de se placer dans une posture d'échange avec les groupes culturels concernés, pour recommencer le processus d'apprentissage culturel, mais aussi pour construire les relations de confiance obligatoires à la bonne tenue de l'évaluation.

L'évaluation indigène et la volonté de décoloniser l'évaluation prolongent et élargissent l'évaluation attentive aux différences culturelles. Dans ce cas, c'est plus largement la « connaissance occidentale » qui est mise à distance; par exemple, les hypothèses relatives à l'individu autonome et rationnel; le dualisme cartésien entre corps et esprit; l'importance donnée à la maîtrise technique au détriment des savoirs d'usage ou de la sagesse pratique, etc. (Wehipeihana et McKegg, 2018). Un enjeu est alors d'imaginer des évaluations s'inscrivant beaucoup plus largement dans des systèmes de connaissance indigènes conçues et mises en œuvre par et avec des évaluateurs et évaluatrices indigènes. C'est ce que recherche notamment le mouvement *Made in Africa* (Frehiwot, 2019). Nan Wehipeihana décrit quant à elle l'évaluation autochtone comme « une évaluation par des personnes autochtones, pour les personnes autochtones, avec des autochtones et en tant qu'autochtones; et où il n'y a pas de rôle tenu pour les non-autochtones, sauf sur invitation » (Wehipeihana, 2019).

### Exemple II – Évaluation autochtone

Il s'agit d'étudier la participation des Maoris à un programme sportif. Un critère classique consisterait à compter le nombre de Maoris dans ce programme. Mais l'équipe d'évaluation cherche à développer des critères prenant en compte la participation « en tant qu'autochtones » : Y avait-il une gouvernance, une gestion, un personnel maori? Parlait-on maori? La mise en œuvre intégrait-elle des pratiques maories? Les activités sportives avaient-elles lieu dans des endroits ayant une signification culturelle pour les Maoris? Les activités sportives elles-mêmes étaient-elles maories? Ces critères sont établis par un processus de dialogue et de délibération avec les opérateurs maoris du programme, et servent à s'accorder sur ce que signifie participer « en tant que Maori » à des activités sportives et récréatives (Wehipeihana et McKegg, 2018).

Une évaluation est engagée au Sénégal avec de jeunes évaluatrices et évaluateurs. Ceux-ci engagent deux approches de collecte de données adaptées à la culture et de nature autochtone. La première, dite « méthodologie Ataya », est un groupe de discussion qui tourne autour de la préparation et de la consommation du thé, un moment très important dans la vie quotidienne des Sénégalaises et des Sénégalais. La seconde est une approche de collecte de données dans laquelle les jeunes évaluatrices et évaluateurs recueillent des données qualitatives auprès d'autres jeunes en utilisant les maximes de Kocc Barma, un philosophe cayorien (avant qu'il ne soit « le Sénégal »).<sup>10</sup>

10. Cette expérience nous a été relatée par Thomas G. Archibald, que nous remercions.

### L'évaluation émancipatrice

Cette approche (*Empowerment Evaluation*) peut être définie comme « le recours à des concepts, techniques et constats issus de l'évaluation pour améliorer [les interventions] et favoriser l'autodétermination » (Fetterman et Wandersman, 2007). L'évaluation émancipatrice est nécessairement une démarche participative. Pourtant, malgré son nom, elle n'est pas nécessairement transformationnelle : elle peut être employée dans un but très pratique de renforcement des capacités d'acteurs impliqués dans un projet, ou pour intégrer l'évaluation dans la routine organisationnelle d'une institution, par exemple. Nous choisissons cependant ici de l'aborder sous l'angle de la justice sociale et de sa promesse : redonner du pouvoir d'agir aux parties prenantes. L'évaluation émancipatrice cherche alors à créer les conditions pour qu'un groupe de personnes, qui peuvent être des travailleurs ou travailleuses de terrain, des destinataires d'une intervention, des groupes marginalisés, engagent une transformation psychologique qui va leur permettre de s'approprier les interventions et de se donner les moyens d'en prendre les rênes.

Il existe toute une famille d'évaluations « de type *empowerment* » (Cousins et Whitmore, 1998; Ridde *et al.*, 2005). Dans la version proposée par David Fetterman, l'évaluation émancipatrice démarre par une étape lors de laquelle le groupe impliqué va s'accorder sur les finalités et les missions de l'intervention évaluée. Les participants et participantes font ensuite un état des lieux de la situation actuelle : où en est-on dans la mise en œuvre et les effets de l'intervention? L'approche recommande le recours à une notation de 1 à 10, qui servira de niveau de référence pour le futur; les participants et participantes doivent engager une négociation pour s'accorder sur cette note, et utiliser des données (ou faire appel à des faits vérifiables) pour appuyer leur opinion. Cette étape peut se répéter à différents stades de l'avancement du programme, permettant ainsi de s'entendre sur l'évolution du programme. La troisième étape donne aux participants et aux participantes la possibilité de se projeter dans le futur, en établissant de nouveaux objectifs, des stratégies pour les atteindre, et les informations à collecter pour les vérifier (Bélanger, 2011).

Dans l'évaluation émancipatrice, l'évaluatrice ou l'évaluateur est un « ami critique ». Elle ou il accompagne, soutient, forme selon les besoins, mais ne détermine pas le processus. Surtout, l'équipe d'évaluation n'organise ni ne mène les investigations, les participants et participantes étant amenés à décider totalement ce dont elles ou ils ont besoin pour leur initiative, ce qui constitue des indices acceptables de la réussite ou de l'échec des interventions.



### **Exemple 12 – Évaluation de type émancipatrice**

Une expérimentation est engagée en France auprès de l'Inspection du travail. Celle-ci vise d'une part à changer les pratiques des inspectrices et des inspecteurs, qui d'ordinaire travaillent seuls, pour passer à une pratique collective; et d'autre part à les amener à évaluer leurs actions. Pour ce faire, le ministère étudie initialement la mise en place d'indicateurs quantitatifs. Cependant, pas question d'envisager une évaluation extérieure de l'Inspection du travail, dont l'autonomie est protégée par une convention internationale, et qui voit dans l'évaluation une tentative de leur hiérarchie de s'immiscer dans leurs pratiques. La solution est de concevoir une démarche mettant les membres de l'Inspection en situation de construire et de mettre en œuvre leurs propres évaluations. Ils sont accompagnés pour cela par un évaluateur qui les forme, les accompagne collectivement dans l'élaboration de leur théorie du changement, de leurs questions d'évaluation et de leurs outils de collecte. Les évaluations proprement dites, cependant, sont menées de bout en bout par les membres de l'Inspection. Outre les apports de la démarche en matière de réflexion collective sur les actions engagées, l'évaluation offre la possibilité aux personnes impliquées de se pencher sur leurs pratiques professionnelles; elle leur donne aussi des compétences et une expérience leur permettant de s'affirmer face aux directives du ministère en matière d'évaluation (Delahais, 2021).

## **Déterminer l'approche qui convient : quels critères de choix ?**

Nous avons dit précédemment que, pour toute évaluation, il fallait déterminer une démarche évaluative ad hoc, et que celle-ci pouvait être élaborée en s'appuyant sur une ou plusieurs approches, elles-mêmes adaptées au contexte d'intervention. Une bonne démarche évaluative dépend ainsi tout autant de l'intervention évaluée, de ce qui est attendu de l'évaluation ainsi que du contexte et des conditions de l'évaluation, mais aussi de ce que les parties prenantes impliquées jugent comme étant une bonne stratégie. Faut-il dès lors redire qu'il n'y a pas une hiérarchie des approches gravée dans le marbre? Plutôt, nous considérerons que toute approche évaluative a des caractéristiques qui lui sont propres, qui lui permettent de briller dans certains contextes ou d'apparaître foncièrement inadaptée dans d'autres.

Pour élaborer une démarche évaluative adaptée aux enjeux, faisable et utile, on pourra ainsi se poser les questions rassemblées dans le tableau 2.



**Tableau 2 — Questions clés pour établir une bonne démarche évaluative**

L'approche est-elle adaptée à l'intervention évaluée?	Quelle est la nature des instruments évalués? L'intervention est-elle innovante, expérimentale, ou au contraire éprouvée? Quels sont les publics concernés, leur nombre, le contexte d'intervention, le type de conséquences attendues et les chemins par lesquels ces impacts peuvent être obtenus, etc.
L'approche répond-elle adéquatement aux questions d'évaluation?	Les questionnements relèvent-ils de l'impact, ou d'autres registres (efficacité, pertinence, cohérence...)? L'approche doit-elle permettre l'émergence des questionnements?
L'approche fournit-elle le type de processus ou de résultats qui rendra l'évaluation utile?	Les attentes portent-elles sur des éléments de description ou d'explication causale? Des conclusions ou des recommandations sont-elles attendues? Y a-t-il des attentes en termes d'inclusion, de participation, d'expertise à prendre en compte?
L'approche peut-elle être crédible dans le contexte où elle sera mise en œuvre? (Hurteau <i>et al.</i> , 2012) <sup>11</sup>	Les commanditaires et les parties prenantes partagent-elles les valeurs exprimées dans l'approche envisagée? S'accordent-elles sur la façon dont l'évaluation devrait être menée, ce à quoi elle devrait servir? Qu'est-ce qui apparaît dans ce cadre comme une preuve crédible?
L'approche est-elle faisable en contexte?	Peut-elle être menée dans les contraintes de temps, de budget et de données disponibles qui sont propres à chaque évaluation (Bamberger <i>et al.</i> , 2021)?
L'équipe d'évaluation est-elle en mesure de mener cette approche?	L'équipe d'évaluation se sent-elle en mesure de la mener à bien? Fait-elle partie de sa zone de confort (en termes de valeurs, de rapport à la connaissance, d'éthique...)? A-t-elle les compétences techniques, sociales et culturelles nécessaires?

On ajoutera que tout cela définit une bonne démarche évaluative a priori, car en réalité, les équipes d'évaluation, leurs commanditaires, voire les autres parties prenantes font des choix tout au long d'une évaluation pour la mener à bien, pour la rendre plus utile, pour s'adapter à des circonstances existantes... sa qualité effective ne peut ainsi être jugée qu'a posteriori. En effet, il est difficile d'imaginer exécuter *fidèlement* une approche dans toutes ses composantes, dans la mesure où celles-ci sont mouvantes. Certaines approches comportent une dimension méthodologique très prescriptive, dont le suivi méticuleux est essentiel pour la validité des résultats obtenus. C'est par exemple le cas pour l'essai contrôlé randomisé, qui fournit des règles strictes sur la composition d'un échantillon ou la tenue d'une expérimentation. Même là, cependant, les marges de manœuvre sont importantes (Jatteau, 2016), et il est utile de conserver une attitude réflexive sur les choix effectués pour s'assurer que l'on fait les meilleurs choix pour l'évaluation. Ainsi, selon le contexte, la capacité d'adaptation d'une stratégie évaluative sera également un critère de qualité à prendre en compte. Dans tous les cas, les équipes d'évaluation seront amenées à faire des choix et pourront s'appuyer sur les approches, surtout lorsque celles-ci expriment clairement des principes, une éthique ou des valeurs pour les guider.

Parce qu'une bonne démarche évaluative dépend d'autant de critères, les choix d'approches qui peuvent être effectués pour l'alimenter sont en réalité limités. De nombreuses personnes préfèrent s'abstraire de ces questions en ne s'appuyant que sur une seule approche, ou sur un petit nombre de méthodes, qui coïncident avec leur croyance en ce qu'une évaluation devrait être, et plus globalement avec leur zone de confort. Des équipes d'évaluation vont ainsi cadrer une évaluation d'une façon qui

11. Toutes les approches que nous proposons ici s'appuient sur des méthodologies éprouvées, mais celles-ci ne sont qu'un élément de la crédibilité (Hurteau *et al.*, 2012).

s'inscrit dans l'approche qu'elles préfèrent, ou ne répondre qu'à des commandes qui s'y inscrivent. Des commanditaires posent leurs questions d'évaluation en fonction de ce à quoi pourra répondre l'approche qui leur est la plus familière, ou qui leur semble la plus crédible.

Dans l'ensemble, gageons que cela ne prêle pas réellement à conséquence. Néanmoins, il est utile de se donner, en début d'évaluation, la possibilité de sortir de la routine en étudiant comment plusieurs approches évaluatives pourraient s'appliquer à une évaluation, et définir ainsi un « champ des possibles ». Cet éventail d'approches pouvant être déployées permet à la fois de mieux répondre à une commande évaluative; mais aussi de réinterroger cette commande, en offrant d'autres façons de mener l'évaluation que celles que l'on envisageait initialement. L'objectif n'est pas forcément d'aboutir à l'approche idéale, mais au meilleur compromis, fait consciemment et dans un cadre précis, pour réaliser une évaluation.

## Conclusion : élargir sa zone de confort

Nous avons, dans les pages précédentes, passé en revue 14 approches et familles d'approches. Certaines sont très semblables à une méthodologie; d'autres touchent à la posture et à la relation au monde; toutes, car c'était un élément essentiel de choix ici, sont utilisées pratiquement par des équipes d'évaluation pour élaborer leurs stratégies évaluatives et guider leurs efforts au quotidien. Cette liste est loin d'être définitive : certaines tomberont peut-être en désuétude dans les prochaines années; d'autres propositions sont peut-être en voie de devenir des approches à part entière.

Chacun est-il amené, dans sa carrière, à utiliser toutes ces approches? Sans doute non : nous avons toutes et tous une « zone de confort », délimitée par nos valeurs, nos habitudes, ce que nous percevons de nos compétences et du champ des possibles en matière d'évaluation. Mais il est possible d'élargir cette zone de confort, en repérant d'autres approches que celles que nous utilisons habituellement; de se redonner une certaine liberté de choix, en considérant plus systématiquement plusieurs stratégies évaluatives; d'hybrider et de sophistiquer nos pratiques, s'inspirer d'autres réflexions et d'autres pratiques que les siennes. C'est ce à quoi ce chapitre aspire.

## Messages clés

- Les approches sont des ensembles de propositions relatives à ce qui constitue une bonne évaluation. Elles combinent de nombreux éléments, et en particulier un cadre paradigmatique, des promesses en termes de résultats et d'usages potentiels, des prescriptions méthodologiques ainsi que des préconisations pratiques permettant de les mettre en œuvre;
- Les approches sont comme des blocs de construction d'une évaluation. S'appuyer sur des approches lors de l'élaboration d'une démarche évaluative ouvre le champ des possibles en matière d'évaluation, permet d'apporter de la crédibilité méthodologique et de mieux se comprendre entre parties prenantes de l'évaluation;
- Il existe de nombreuses approches, qui correspondent à des contextes et des attentes différentes, mais aussi à la vision du monde et aux valeurs des différentes parties prenantes, évaluateurs et évaluatrices y compris. Il faut donc choisir en conséquence celles qui conviennent.

## Références bibliographiques

- Abma, T. A. (1998). Storytelling as Inquiry in a Mental Hospital. *Qualitative Health Research*, 8(6), 821-838. <https://doi.org/10.1177/104973239800800608>
- Alkin, M. C. (2013). *Evaluation roots : A wider perspective of theorists' views and influences* (2<sup>e</sup> éd.). SAGE Publications.
- Bamberger, M., Rugh, J., Church, M. et Fort, L. (2021). L'évaluation en situation réelle : concevoir des évaluations d'impact sous contraintes de budget, de temps et de données (trad., publication originale en 2004). Dans T. Delahais, A. Devaux-Spatarakis, A. Revillard, et V. Ridde (dir.), *Évaluation. Fondements, Controverses et Perspectives*. Éditions science et bien commun. <https://scienceetbiencommun.pressbooks.pub/evaluationanthologie/>
- Barret, D., Blundo Canto, G., Dabat, M.-H., Devaux-Spatarakis, A., Faure, G., Hainzelin, E., Mathé, S., Temple, L., Toillier, A. et Triomphe, B. (2017). *Guide méthodologique ImpresS. Évaluation ex post des impacts de la recherche agronomique dans les pays du Sud*. Cirad.
- Behaghel, L., Crépon, B., Gurgand, M., Kamionka, T., Lequien, L., Rathelot, R. et Zamora, P. (2013). L'accompagnement personnalisé des demandeurs d'emploi. *Revue française d'économie*, 28(1), 123-158. <https://doi.org/10.3917/rfe.131.0123>
- Befani, B. et Mayne, J. (2014). Process Tracing and Contribution Analysis : A Combined Approach to Generative Causal Inference for Impact Evaluation. *IDS Bulletin*, 45(6), 17-36. <https://doi.org/10.1111/1759-5436.12110>
- Befani, B. (2020). *Choosing Appropriate Evaluation Methods Tool* (version 2). CECAN.
- Belcher, B. M., Davel, R. et Claus, R. (2020). A refined method for theory-based evaluation of the societal impacts of research. *MethodsX*, 7. <http://dx.doi.org/10.1016/j.mex.2020.100788>
- Bélangier, J.-F. (2011). L'évaluation habilitative. Dans P. M. Daigneault (dir.), *Les approches théoriques en évaluation*. Perfeval.
- Brousselle, A. et Buregeya, J.-M. (2018). Theory-based evaluations : Framing the existence of a new theory in evaluation and the rise of the 5th generation. *Evaluation*, 24(2), 153-168. <https://doi.org/10.1177/1356389018765487>
- Chen, H. T. et Rossi, P. H. (1987). The theory-driven approach to validity. *Evaluation and Program Planning*, 10(1), 95-103. [https://doi.org/10.1016/0149-7189\(87\)90025-5](https://doi.org/10.1016/0149-7189(87)90025-5)
- Copestake, J. G., Morsink, M. et Remnant, F. (2019). *Attributing development impact : The Qualitative Impact Protocol (QuIP) case book*. Practical Action Publishing.
- Cousins, J. B. et Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 1998(80), 5-23. <https://doi.org/10.1002/ev.1114>
- Dagenais, C. et Ridde, V. (2012). *Approches et pratiques en évaluation de programmes*. Les Presses de l'Université de Montréal.
- Delahais, T. et Toulemonde, J. (2017). Making rigorous causal claims in a real-life context : Has research contributed to sustainable forest management? *Evaluation*, 23(4), 370-388. <https://doi.org/10.1177/1356389017733211>

- Delahais, T., Sage, K. et Honoré, V. (2020). Evaluators in Transition. *Zeitschrift für Evaluation (ZfEv)*, 2020(2), 239-260. DOI :10.31244/zfe.2020.02.03
- Delahais, T. (2021). Les évaluations participatives : Plus facile à dire qu'à faire? *LeGes*, 32(1). <https://doi.org/10.38023/265714d6-463d-4ae3-80ef-5588489d0ec3>
- deMarrais, K. B et Lapan, S. D. (dir.). (2004). *Foundations for research : Methods of inquiry in education and the social sciences*. Routledge –Taylor & Francis Group.
- Demarteau, M. (2002). A theoretical framework and grid for analysis of programme-evaluation practices. *Evaluation*, 8(4), 454-473. <https://doi.org/10.1177/13563890260620649>
- Devaux-Spatarakis, A., Harmach, S., Alouis, J. et Sage, K. (2021). Concilier les différents intérêts d'apprentissage d'une expérimentation : retour d'expérience sur le rôle de l'évaluateur pour favoriser les usages de l'évaluation. Dans A. Kerivel et S. James (dir.), *Comment évaluer des projets innovants encore en construction?* INJEP Notes & Rapports 2021/14.
- Dozois, E., Langlois, M. et Blanchet-Cohen, N. (2011). *DE 201 : Guide du praticien de l'évaluation évolutive*. Fondation de la famille J.W. McConnell et l'Institut international des droits de l'enfant et du développement.
- Earl, S., Carden, F. et Smutylo, T. (2001). *Outcome Mapping : Building Learning and Reflection into Development Programs*. IDRC.
- Fetterman, D. et Wandersman, A. (2007). Empowerment Evaluation : Yesterday, Today, and Tomorrow. *American Journal of Evaluation*, 28(2), 179-198. <https://doi.org/10.1177/1098214007301350>
- Frehiwot, M. (2019). *Made in Africa Evaluation : Decolonizing evaluation in Africa* (third quarter 2019). EvalMatters. <http://idev.afdb.org/sites/default/files/Evaluations/2020-03/article%203%20Decolonizing%20evaluation%20in%20Africa.pdf>
- Funnell, S. C. et Rogers, P. J. (2011). *Purposeful program theory : effective use of theories of change and logic models*. Jossey-Bass.
- Gibson, R. et Robichaud, S. (2020). Evaluating Dancing with Parkinson's : Reflections from the perspective of a community organization. *Evaluation and Program Planning*, 80(101449). <https://doi.org/10.1016/j.evalprogplan.2017.05.010>
- Guba, E. G. et Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. SAGE Publications.
- Harji, K. et Jackson, E. T. (2016). *Developmental evaluation in practice : Lessons from evaluating a market-based employment initiative*. The Rockefeller Foundation. <https://www.issuelab.org/resources/27188/27188.pdf>
- Hood, S., Hopson, R. K. et Kirkhart, K. E. (2015). Culturally responsive evaluation. Dans E. K. Newcomer, P. H. Hatry et S. J. Wholey (dir.), *Handbook of practical program evaluation*. Wiley. <https://doi.org/10.1002/9781119171386.ch12>
- House, E. et Howe, K. (2003). Deliberative Democratic Evaluation. Dans T. Kellaghan et D. L. Stufflebeam (dir.), *International Handbook of Educational Evaluation* (p. 79-100). Springer.
- Hurteau, M., Houle, S. et Guillemette, F. (dir.). (2012). *L'évaluation de programme axée sur le jugement crédible*. Presses de l'Université du Québec.

- Jatteau, A. (2020). *Faire preuve par le chiffre? Le cas des expérimentations aléatoires en économie*. Institut de la gestion publique et du développement économique.
- Lay, M. et Papadopoulos, I. (2007). An Exploration of Fourth Generation Evaluation in Practice. *Evaluation*, 13(4), 495-504. <https://doi.org/10.1177/1356389007082135>
- Mayne, J. (2012). Contribution analysis : Coming of age? *Evaluation*, 18(3), 270-280. <https://doi.org/10.1177/1356389012451663>
- Mertens, D.M. et Wilson, A.T. (2012). *Program Evaluation Theory and Practice : A Comprehensive Guide*. Guilford Press.
- Mertens, D. (2021). La recherche transformationnelle : personnelle et sociétale (trad., publication originale en 2017). Dans T. Delahais, A. Devaux-Spatarakis, A. Revillard, et V. Ridde (dir.), *Évaluation. Fondements, Controverses et Perspectives*. Éditions science et bien commun. <https://scienceetbiencommun.pressbooks.pub/evaluationanthologie/>
- Miller, R.L. (2010). Developing Standards for Empirical Examinations of Evaluation Theory. *American Journal of Evaluation*, 31(3), 390-399. <https://doi.org/10.1177/1098214010371819>
- ONU Femmes (2015). *Gestion des évaluations sensibles au genre*. <https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/Library/Publications/2015/UN-Women-Evaluation-Handbook-fr.pdf>
- Patton, M.Q. (2008). *Utilization-focused evaluation* (4<sup>e</sup> éd.). SAGE Publications.
- Patton, M.Q. et LaBossière, F. (2012). L'évaluation axée sur l'utilisation. Dans C. Dagenais et V. Ridde (dir.), *Approches et pratiques en évaluation de programmes* (p. 145-160). Presses de l'Université de Montréal.
- Pawson, R. et Tilley, N. (1997). *Realistic evaluation*. SAGE Publications.
- Podems, D. (2014). Feminist Evaluation for non-feminists. Dans S. Brisolaro, D.M. Seigart et S. SenGupta (dir.), *Feminist evaluation and research : Theory and practice*. The Guilford Press.
- Podems, D. (2018). *Rendre l'évaluation féministe praticable* (4<sup>e</sup> trimestre, 2018). Evaluation Matters. <https://idev.afdb.org/sites/default/files/Evaluations/2020-03/Rendre%20l%27%20evaluation%20feministe%20praticable.pdf>
- Quadrant Conseil. (2017). *Comment évaluer l'impact?* <https://www.quadrant-conseil.fr/ressources/ArbreImpact.html>
- Ridde, V., Baillargeon, J., Ouellet, P. et Roy, S. (2005). L'évaluation participative de type empowerment : Une stratégie pour le travail de rue. *Service social*, 50(1), 263-279. <https://doi.org/10.7202/009972ar>
- Robert, É. et Ridde, V. (2013). L'approche réaliste pour l'évaluation de programmes et la revue systématique : De la théorie à la pratique. *Mesure et évaluation en éducation*, 36(3), 79-108. <https://doi.org/10.7202/1025741ar>
- Shadish, W. R., Cook, T. D. et Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Company.

- Sielbeck-Mathes, K. et Selove, R. (2014). An Explication of Evaluator Values : Framing Matters. Dans S. Brisolaro, D. Seigart et S. SenGupta (dir.), *Feminist evaluation and research : Theory and practice*. Guilford Publications.
- Stake, R. (2003). Responsive Evaluation. Dans T. Kellaghan et D.L. Stufflebeam (dir.), *International Handbook of Educational Evaluation* (p. 63-68). Springer.
- Stufflebeam, D. L. et Coryn, C.L.S. (2012). *Evaluation Theory, Models, and Applications* (2<sup>e</sup> éd.). Jossey-Bass.
- Turcotte, P.-L., Rey, L. et Brousselle, A. (2021). L'évaluation évolutive, de la théorie à la pratique : Perspectives de praticiens québécois. *Canadian Journal of Program Evaluation*, 36(1). <https://doi.org/10.3138/cjpe.70502>
- Vaessen, J., Lemire, S. et Befani, B. (2020). *Evaluation of International Development Interventions : An Overview of Approaches and Methods*. World Bank.
- Wehipeihana, N. et McKegg, K. (2018). Values and Culture in Evaluative Thinking : Insights from Aotearoa New Zealand : Values and Culture in Evaluative Thinking. *New Directions for Evaluation*, 2018(158), 93-107. <https://doi.org/10.1002/ev.20320>
- Wehipeihana, N. (2019). Accroître la compétence culturelle, une étape nécessaire en appui à une évaluation menée par les personnes autochtones. Dans T. Delahais, A. Devaux-Spatarakis, A. Revillard et V. Ridde (dir.), *Évaluation. Fondements, Controverses et Perspectives*. Éditions Sciences du Bien Commun.
- Wenger-Trayner, E. et Wenger-Trayner, B. (2015, 15 avril). *Communities of practice : A brief introduction*. Wenger-Trayner. <https://wenger-trayner.com/wp-content/uploads/2015/04/07-Brief-introduction-to-communities-of-practice.pdf>
- White, H. (2009). *L'évaluation d'impact basée sur la théorie : Principes et pratique*. Wageningen University & Research.
- White, H. (2014). *Essais contrôlés randomisés (ECR)* (Note méthodologique n° 7). UNICEF – Centre de recherche Innocenti.
- White, H. et Sabarwal, S. (2014). *Méthodes et modèles quasi expérimentaux* (Note méthodologique n° 8). UNICEF – Centre de recherche Innocenti.
- Wilson-Grau, R. et Britt, H. (2012). *Outcome harvesting*. Ford Foundation.



# Les approches économiques en évaluation : un aperçu

*Eva Anstett  
Miché Ouédraogo*

## Introduction

Plusieurs approches sont utilisées en évaluation des politiques publiques, parmi lesquelles on retrouve l'évaluation économique. Cette dernière est un processus qui consiste à comparer les bénéfices et les coûts respectifs de différentes interventions. L'objectif de cette approche est en effet de mesurer l'efficacité, ou l'utilité de l'argent dépensé pour une intervention (Boardman et Vining, 2017; Goldsmith *et al.*, 2004). Ainsi, pour chaque intervention, on considère le coût d'opportunité, c'est-à-dire qu'on se demande si la somme consacrée à celle-ci aurait pu être utilisée à une autre intervention ayant potentiellement un rendement plus élevé, ou plus simplement si on aurait mieux fait d'investir l'argent public dans une autre intervention. Dans un contexte de rareté de ressources où l'on veut faire mieux avec moins, il est essentiel d'évaluer économiquement les différentes options. Ainsi, lorsque les décideurs publics doivent prendre une décision quant à une intervention, l'évaluation économique s'avère être un outil qui permet de produire et de représenter des données sur les plans économique, social, politique, etc., de proposer des alternatives et de faire ainsi des choix judicieux en faisant un bon usage des ressources publiques (Wilkinson *et al.*, 2016). Il existe plusieurs méthodes d'évaluation économique des interventions. Il s'agit de l'analyse coût-bénéfice (ACB), l'analyse coût-efficacité (ACE), l'analyse coût-utilité (ACU) et l'analyse coût-conséquence (ACC). Ce chapitre présente succinctement ces méthodes ainsi que les enjeux de l'évaluation économique.

## Présentation des principales méthodes d'évaluation économique des interventions

L'**analyse coûts-bénéfices (ACB)** ou analyse avantage-coût tire ses racines dans la théorie économique du welfarisme, et repose sur le principe de la réalisation de toute décision pour laquelle les bénéfices sont supérieurs aux coûts, et ceci pour l'ensemble de la société (CCNPPS, 2014). Elle recherche la maximisation des gains nets en matière de bien-être collectif. Cette méthode représente tous les coûts et tous les bénéfices (positifs ou négatifs) des interventions en valeur monétaire, en utilisant par exemple les préférences révélées ou la volonté de payer, afin de mettre un prix sur des bénéfices n'ayant pas de valeur marchande bien déterminée (Boardman *et al.*, 1994). Elle repose donc principalement sur l'hypothèse voulant que non seulement tous les coûts, mais aussi tous les effets d'une intervention puissent être exprimés en unités monétaires. Une fois les coûts et les bénéfices déterminés, la recommandation consiste à choisir l'intervention qui a le ratio coûts-bénéfices ou la valeur actuelle nette (VAN) les plus élevés (Meunier et Marsden, 2009).

Si l'on ne veut pas se lancer dans une mesure monétaire de chaque coût et bénéfice, on peut utiliser l'**analyse coût-efficacité (ACE)** qui permet d'exprimer les bénéfices d'une intervention non pas en valeur monétaire, mais dans des mesures appropriées au problème à résoudre et à l'intervention. Elle

réduit par contre les possibilités de comparaison, car les interventions doivent viser une même variable et doivent rester comparables sur cette base unique (Donaldson et Shackley, 2003; Vining et Boardman, 2004). Ainsi, les différentes options ou alternatives d'intervention disponibles sont classées selon un ratio de coût pour obtenir un gain X, où X peut être n'importe quelle unité de mesure naturelle (p. ex. : nombre de personnes vaccinées, km de pistes cyclables, etc.) (CCNPPS, 2014).

L'**analyse coût-utilité (ACU)** va un peu plus loin que l'ACE et analyse les différentes interventions alternatives, non pas à l'aide d'un seul critère (km de pistes cyclables), mais en se basant sur un indicateur composite qui représente une mesure plus globale des effets d'une intervention (Wagstaff, 1991). Pour cela, l'ACU utilise souvent dans le secteur de la santé une mesure de qualité de la vie : l'année de vie ajustée en fonction de la qualité, mieux connue sous l'acronyme AVAQ (ou QALY en anglais) (Briggs, 1999; Vining et Boardman, 2004; Weinstein *et al.*, 2009). Les différentes interventions possibles sont classées selon leur coût pour chaque AVAQ gagnée, l'AVAQ variant entre 0 et 1.

Enfin, l'**analyse coût-conséquence (ACC)** est une approche qui consiste à combiner les différentes approches, ACB, ACE et ACU, dans une seule évaluation d'efficacité. Ainsi, les différentes options d'intervention possibles sont classées dans un tableau, et les résultats des différentes méthodes sont placés dans les colonnes. Ce tableau-synthèse permet aux décideurs d'avoir une vue générale de l'efficacité des différentes options.

Maintenant que les bases des approches économiques sont présentées, il est intéressant de se pencher sur leurs avantages et leurs limites.

## Enjeux, avantages et limites des approches économiques

Comme toute approche d'évaluation, les approches économiques ne s'appliquent pas toujours de la même façon ni dans toutes les situations.

### Une méthode de calcul standardisée pour répondre aux questions d'efficacité

Un des avantages évidents des approches économiques est leur méthode de calcul standardisée : on utilise des définitions communément admises dans la communauté des évaluateurs (ce qui est rarement le cas pour d'autres approches), des formules de calculs et des critères de décision généralement clairs. En effet, si la VAN d'un programme à l'étude est positive, alors on peut affirmer que celui-ci est rentable, et qu'il s'agirait d'un bon investissement pour la société (ACMITS, 2017). Cette simplicité apparente ne veut toutefois pas dire que les évaluateurs se transforment en techniciens, qui appliquent juste des formules toutes faites. Une réflexion sur le type d'analyse à mettre en œuvre, sur les hypothèses, sur les coûts et avantages à prendre en compte, sur les marges d'erreur, voire sur la façon de présenter les chiffres, reste indispensable. De plus, même si l'évaluation économique est bien réalisée, elle ne se rapporte qu'au critère de l'efficacité, il est donc utile de confronter ses résultats à d'autres investigations, notamment sur l'équité, qui est absente de la réflexion économique (un avantage reste un avantage, peu importe qui en est le bénéficiaire) et qui est pourtant centrale pour des interventions visant le développement (CCNPPS, 2014). En outre, le fait d'ajouter l'ensemble des effets d'une intervention dans la balance (par exemple pour le calcul de la VAN) nécessite qu'une évaluation des effets soit réalisée. Les questions d'évaluations sur le critère d'efficacité sont donc liées aux réponses et méthodes appliquées pour répondre aux questions liées aux effets (ou critère d'impact).



## Effectuer des comparaisons

Un autre intérêt principal des approches économiques est leur capacité à placer des interventions différentes vis-à-vis d'une métrique commune et donc de permettre leur comparaison : plusieurs scénarios d'action (évaluation *ex ante*); forces et faiblesses de différentes mises en œuvre (différents sites, différentes régions); ou encore, faire le bilan d'interventions de nature différente, mais visant une même cible ultime (par exemple un même objectif de développement durable (ODD)). Néanmoins, comme pour toute comparaison, il ne faut pas négliger le fait que ces chiffres nous renseignent sur l'efficacité d'une intervention sans en donner une image complète. Comparer les ratios coût-efficacité de deux régions alors qu'elles desservent des populations différentes (une plus urbaine, une autre rurale par exemple) peut donner aux lecteurs une fausse impression si le contexte de la différence de ratio n'est pas bien expliqué (Pierce *et al.*, 2006). Encore une fois, il revient aux évaluateurs de ne pas utiliser ces méthodes à tort et à travers, et de ne pas présenter de chiffres sans réfléchir à leur pertinence dans un argumentaire évaluatif crédible.

## L'enjeu de la valeur monétaire

Au-delà des méthodes de calculs standardisées, il n'est jamais évident de donner une valeur monétaire à des coûts ou à des avantages sociaux ou environnementaux. Comment chiffrer la valeur d'une forêt primaire protégée? D'une tonne de gaz à effet de serre rejetée? Du bouleversement du mode de vie d'une communauté par la désertification? Pourtant, c'est bien ce type de variables qui forment des cibles souhaitées des interventions de développement et qui doivent être prises en compte dans une évaluation économique exhaustive. De nombreuses méthodes ont été développées dans la littérature pour donner des estimations à ces effets souvent intangibles, mais aucune d'entre elles n'est parfaite et ne s'applique dans toutes les situations (Meunier et Marsden, 2009; Pierce *et al.*, 2006). Une réflexion éthique est également à mener face à ces méthodes qui nécessitent de donner une valeur à tous les effets, du bien-être des populations, à la biodiversité en passant par les vies humaines. Il apparaît en revanche intéressant, justement dans le cadre d'interventions visant le développement (complexes, avec des impacts à long terme et avec des visées humaines et sociales), d'explicitier la valeur des impacts recherchés pour les mettre en regard de coûts qui paraissent souvent trop élevés au moment de la dépense, mais qui seraient finalement moins élevés que les conséquences négatives du *statu quo* (CCNPPS, 2014). Dans ce sens, dans un monde de ressources rares, où les interventions même publiques sont souvent vues par le prisme de leurs dépenses, les approches économiques peuvent être la base d'un plaidoyer argumenté pour des interventions de développement positives pour leurs sociétés.

## Conclusion

Les approches économiques d'évaluation sont donc des outils intéressants pour répondre aux questions d'efficacité des interventions, notamment d'un point de vue comparatif entre différentes variantes d'interventions, voire de plusieurs interventions. Bien que leur apparition ne soit pas récente dans le secteur public, la réflexion sur le rapport entre les dépenses publiques et leurs résultats est on ne peut plus actuelle, ce qui rend les méthodes économiques encore pertinentes, et de nombreux évaluateurs continuent d'ailleurs de proposer des résultats et de nouvelles méthodes intéressantes. Leur utilisation n'est toutefois pas sans enjeux, du fait que le principe même des méthodes vise à envisager les interventions sous l'angle monétaire, et qu'elles sont techniquement compliquées. Elles restent toutefois un outil intéressant pour les évaluations *ex ante* de dépenses très importantes, où le temps mis dans l'analyse se justifie, ou pour donner une valeur à des coûts et avantages intangibles, pour lesquels un plaidoyer de financement est parfois difficile à établir sans argument monétaire.

## Pour aller plus loin

Ce guide fait partie du peu de ressources disponibles en français sur les approches économiques :

Meunier, V. et Marsden, É. (2009). *Analyse coût-bénéfices : guide méthodologique* : FonCSI.

Ce livre utilise l'approche économique pour évaluer une nature d'intervention en particulier, soit les réglementations :

Lamari, M., Anstett, E. et Bouchard, J. (2015) *Analyse d'impact réglementaire (AIR) : balises méthodologiques pour mieux évaluer les réglementations*. Presses de l'Université du Québec.

## Références bibliographiques

ACMTS. (2017). *Lignes directrices de l'évaluation économique des technologies de la santé au Canada*, 4<sup>e</sup> éd. Ottawa.

Boardman, A. E., Mallery, W. L. et Vining, A.R. (1994). Learning from ex ante/ex post cost-benefit comparisons : the Coquihalla highway example. *Socio-Economic Planning Sciences*, 28(2), 69-84.

Boardman, A. E. et Vining, A.R. (2017). There are many (well, more than one) paths to Nirvana : The economic evaluation of social policies. Dans B. Greeve (dir.), *Handbook of Social Policy Evaluation*. Edward Elgar Publishing.

Briggs, A. (1999). Handling uncertainty in economic evaluation. *BMJ*, 319(7202), 120.

CCNPPS. (2014). *Méthodes d'évaluation économique : Quelles sont les implications éthiques pour les politiques publiques favorables à la santé?*

Donaldson, C. et Shackley, P. (2003). Willingness to pay for health care. *Advances in health economics*, 1-24.

Goldsmith, L., Hurley, J. E. et Hutchison, B.G. (2004). *Economic evaluation across the four faces of prevention : a Canadian perspective*. Citeseer.

Meunier, V. et Marsden, É. (2009). *Analyse coût-bénéfices : guide méthodologique*. FonCSI.

Pierce, D., Atkinson, G. et Mourato, S. (2006). *Analyse coûts-bénéfices et environnement, développements récents*. OCDE.

Vining, A.R. et Boardman, A.E. (2004). *Policy analysis methods in Canada*. Paper presented at the Policy Analysis Workshop at Simon Fraser University.

Wagstaff, A. (1991). QALYs and the equity-efficiency trade-off. *Journal of health economics*, 10(1), 21-41.

Weinstein, M.C., Torrance, G. et McGuire, A. (2009). QALYs : the basics. *Value in health*, 12, S5-S9.

Wilkinson, T., Sculpher, M. J., Claxton, K., Revill, P., Briggs, A., Cairns, J. A., Teerawattananon, Y., Asfaw, E., Lopert, R., Culyer, A. J. et Walker, D.G. (2016). The international decision support initiative reference case for economic evaluation : an aid to thought. *Value in health*, 19(8), 921-928.

# *Value for money* : Apprécier la valeur d'une intervention basée sur la performance de ses investissements

*Lynda Rey*

*Alexandre Fortin*

## **Introduction**

Pour mesurer le rendement d'une intervention, soit le rapport entre les effets et les ressources financières consacrées à celle-ci (Brousselle *et al.*, 2018, p. 193), de nombreux évaluateurs et évaluatrices peuvent avoir le réflexe d'opter pour des approches économiques d'évaluation. Pourtant, bien que ces dernières comportent des atouts considérables, voire indispensables, elles comportent certaines limites qui devraient être prises en compte par les évaluateurs et évaluatrices. Comme mentionné dans un chapitre précédent, Anstett et Ouédraogo soulignent que « les approches économiques d'évaluation sont [donc] des outils intéressants pour répondre aux questions d'efficience des interventions [...] [mais] leur utilisation n'est toutefois pas sans enjeu, du fait que le principe même des méthodes vise à envisager les interventions sous l'angle monétaire, et qu'elles sont techniquement compliquées ». En effet, leur applicabilité peut varier en fonction du contexte et du type d'intervention en question, plus particulièrement en lien avec la quantification des effets. Par exemple, dépendamment du degré de complexité de l'intervention, l'existence d'une multitude d'effets peut contribuer à accroître le temps et les ressources nécessaires à leur mesure. Aussi, il existe souvent des interactions ou des chevauchements entre différents effets provenant de plusieurs interventions implantées dans le même environnement et qui peuvent rendre difficiles leurs attributions respectives (King et OPM, 2018, p. 17). De plus, l'expertise technique requise liée à la complexité des méthodes utilisées dans le cadre de ces approches peut limiter aussi son utilisation, plus particulièrement pour les petites organisations, notamment celles qui valorisent des effets intangibles tels que l'acquisition de savoir-faire par les participants comme la négociation, l'écoute, la confiance en soi et le travail d'équipe (King, 2019b). La primauté de la valeur de l'efficience à laquelle elle est associée peut constituer aussi un frein à son utilisation, notamment pour les évaluateurs et évaluatrices ne s'inscrivant pas dans le paradigme postpositiviste ou pour des secteurs d'intervention à dimension sociale qui privilégient d'autres valeurs, telles que l'équité et la liberté dans le secteur de la santé (Brousselle *et al.*, 2018). Enfin, les approches économiques en évaluation comportent un certain risque sur le plan méthodologique en lien avec la surestimation de l'importance de certains indicateurs (facilement mesurables, souvent quantitatifs) par rapport à d'autres (plus difficiles à mesurer, souvent qualitatifs) (King, 2017). En effet, la négligence ou l'omission de certaines variables importantes moins évidentes à mesurer dans le cadre de telles approches pourrait se traduire par des observations biaisées et des constats erronés ou incomplets sur le rendement de l'intervention.

En considérant ces limites, il peut être plus judicieux d'opter pour l'utilisation d'un cadre évaluatif plus global, flexible et adapté à des circonstances particulières en allant au-delà de la mesure du rendement d'une intervention pour considérer la performance de celle-ci en tenant compte de l'investissement global que sa mise en œuvre exige. L'approche de la *Value for Money* (VFM) de Julian King (2017) a été développée avec cette intention, notamment pour les interventions complexes comme celles implantées en contexte de développement international. Il s'agit d'une approche qui vise à déterminer le niveau de performance globale d'une intervention à l'aide d'un cadre comportant huit étapes, et dont la ligne directrice est de rendre le plus explicite possible la démonstration et le raisonnement utilisés dans le processus de la collecte et de l'analyse des données (King et OPM, 2018). Cette approche n'ambitionne pas de remplacer les approches traditionnelles en la matière, mais propose de les envisager dans un cadre plus large, permettant d'apprécier à la fois l'investissement et ses conséquences dans le cadre d'une intervention sociale (King, 2019a). Ce chapitre réflexif vise à présenter l'approche VFM qui semble prometteuse, plus particulièrement pour apprécier la performance globale des interventions intégrant une dimension sociale et d'équité. Il a pour objectif d'inciter la communauté évaluative francophone d'Afrique et d'Haïti à s'en inspirer, à l'expérimenter en l'appliquant dans l'élaboration et l'exécution du devis évaluatif d'une intervention particulière, pour laquelle elle pourrait être pertinente.

## L'approche de la VFM de King : une conception plus élargie des approches économiques en évaluation

Plus précisément, qu'est-ce que l'approche *Value for Money* conceptualisée par King? Selon l'auteur, elle n'a pas de définition spécifique. D'abord, elle peut être considérée comme étant « une question évaluative, exigeant un jugement basé sur des arguments logiques et des preuves. Deuxièmement, la VFM représente un domaine commun à deux disciplines, car elle concerne le mérite, la valeur et l'importance d'une intervention (le domaine de l'évaluation) et l'allocation des ressources (le domaine de l'économie) » (traduction libre de King, 2019a, p. 3). En somme, il s'agit d'apprécier « comment les ressources [monétaires et non monétaires] sont utilisées et si elles le sont correctement » (King et OPM, 2018, p. 4). Les définitions opératoires existantes peuvent varier en fonction de l'organisation; elles peuvent mettre l'accent sur différents éléments tels que : « la minimisation du gaspillage, la livraison d'extrants, l'atteinte des résultats, favoriser l'équité et/ou maximiser les effets pour un coût donné, la priorité souvent donnée à cette dernière » (King et OPM, 2018, p. 7). Le concept de la VFM est donc plus englobant dans la mesure où elle peut impliquer la prise en compte de différentes dimensions telles que la qualité, la valeur ou l'équité (King et Guimaraes, 2016), alors que les approches économiques traditionnelles en évaluation sont davantage centrées sur des critères strictement économiques, plus particulièrement l'efficience (Drummond *et al.*, 2005; Fox-Rushby et Cairns, 2005).

À l'instar des autres approches de la VFM (IDEV, 2016; Fleming, 2013; Jackson, 2012a, 2012b; Glendinning, 1988), celle de King mise sur un raisonnement évaluatif explicite, c'est-à-dire que tout jugement posé au cours de la démarche évaluative doit être détaillé et expliqué clairement (King et OPM, 2018, p. 9). L'objectif ici est d'éviter de poser des jugements subjectifs non justifiés, d'assurer une transparence quant au caractère systématique de l'évaluation et de construire un raisonnement robuste et solide. L'approche préconise de bien définir « ce qu'est une bonne VFM » (King et OPM, 2018, p. 4) dans le cadre d'une démarche participative impliquant les parties prenantes dès la conception de l'évaluation.

Pour y parvenir, des critères de la VFM reflétant les dimensions et standards de la performance qui seront importants dans le jugement évaluatif doivent être eux aussi définis en collaboration avec les parties prenantes (King et OPM, 2018, p. 23-24). Ces critères permettent de déterminer le mérite, la valeur et l'importance des ressources utilisées, et devront être sélectionnés de manière pragmatique selon le contexte de l'intervention. Ils peuvent concerner autant les ressources utilisées, les extrants, les résultats et les impacts. King propose de considérer au moins cinq critères, soit « les 4 E » (économie, efficacité, efficacité, équité), en plus du critère de coût-efficacité en vue de comparer les coûts aux effets, ou l'impact aux coûts totaux. Pour chacun de ces **critères**, des définitions spécifiques et directement liées au fonctionnement prévu de l'intervention ainsi qu'à ses effets (en lien avec la théorie de changement et/ou le modèle logique) devront être élaborées. Lors de cette étape, King insiste sur l'importance de ne pas tenir compte du caractère mesurable des critères, car ceux-ci se distinguent des indicateurs : « Alors que les indicateurs sont spécifiques et mesurables, les critères en revanche, décrivent la nature des changements prévus, et ces descriptions sont délibérément plus larges et moins spécifiques, reflétant leur objectif de faciliter un jugement évaluatif qui a plus de sens » (King et OPM, 2018, p. 13). En ce qui concerne les **standards de performance**, King propose de classer les différents niveaux (par exemple : excellent, bon, acceptable, faible), et insiste sur l'importance d'établir des définitions opérationnelles pour chacun d'entre eux, et ce, pour tous les critères. Pris ensemble, les critères et les standards seront présentés sous forme de tableaux (des « rubriques »)<sup>1</sup> (tableau 1) et devront refléter les différents éléments de la théorie de changement (intrants, activités, extrants, résultats et impacts). Finalement, King réitère l'importance d'impliquer les bonnes parties prenantes dans l'élaboration des rubriques, de faire abstraction de tout aspect en lien avec la mesure, d'utiliser des références (*benchmarking*) lorsque c'est possible, de les adapter au contexte de l'intervention évaluée, de viser une cohérence entre les différents standards de performance et de demeurer dans la simplicité dans le choix des critères.

**Tableau 1 — Rubrique relative à la performance d'un projet. Traduit librement par les auteurs et reproduit de King, 2020, p. 12, avec sa permission.**

Standards de performance	Définitions
Excellent	Le projet <b>atteint ou dépasse</b> toutes les attentes/objectifs réalisables en tenant compte du contexte. Possibilité d'amélioration progressive.
Bon	Le projet <b>répond généralement</b> aux attentes/objectifs raisonnables en tenant compte des attentes mineures. Quelques améliorations nécessaires.
Acceptable	Le projet ne répond pas aux attentes/objectifs, <b>mais satisfait aux exigences minimales</b> et montre des <b>progrès acceptables</b> dans l'ensemble. Des améliorations importantes nécessaires.
Faible	<b>Ne satisfait pas aux exigences minimales</b> , aux exigences de base ou ne montre pas de progrès acceptables dans l'ensemble. Améliorations urgentes nécessaires.

1. Pour une explication des rubriques, voir King *et al.* (2013). Evaluative rubrics : A method for surfacing values and improving the credibility of evaluation. *Journal of MultiDisciplinary Evaluation*, 9(21), 11-20.

Ainsi, pour chaque critère (économie, efficacité, équité et coût-efficacité), il convient de développer une rubrique. L'utilisation de critères et de standards permet de répondre directement à la question évaluative relative à l'appréciation globale de la performance d'une intervention au regard des investissements consentis. Le tableau suivant synthétise l'interprétation que l'on fait pour chacun des critères et offre une vue d'ensemble du jugement évaluatif selon les critères et standards convenus; les éléments de preuve qui étayaient ce jugement et les opportunités d'amélioration (tableau 2).

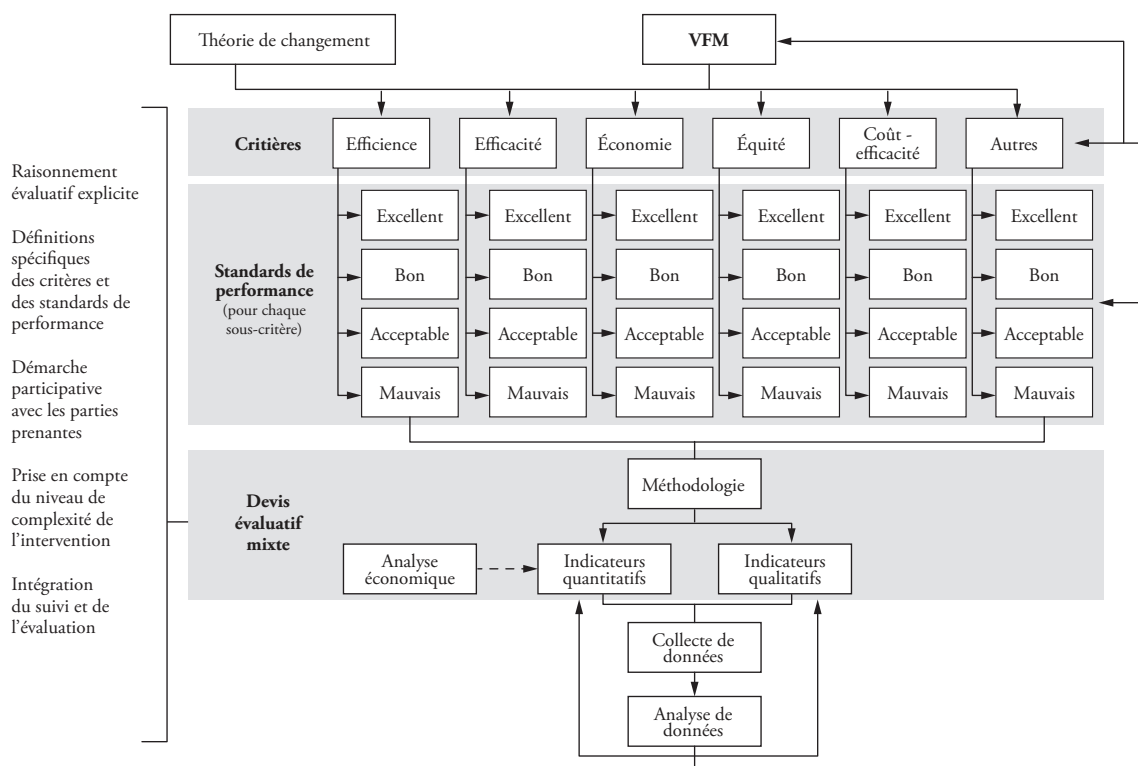
**Tableau 2 — Appréciation de la performance globale de l'intervention. Traduit librement par les auteurs et reproduit de King, 2020, p. 14, avec sa permission.**

Critères	Jugement	Éléments de preuve	Opportunités d'amélioration
Économie	Excellent		
Efficience	Bon		
Efficacité	Bon		
Équité	Acceptable		
Coût-efficacité	Bon		
Performance globale (VFM)	Bon		

L'approche de la VFM de King présuppose l'élaboration d'un devis évaluatif mixte combinant les méthodes quantitatives et qualitatives. La principale prémisse qui supporte cette prescription est associée à une limite des approches économiques en évaluation mentionnée précédemment, soit que les indicateurs quantitatifs ne peuvent pas refléter à eux seuls la performance d'une intervention relativement complexe, et que celle-ci nécessite aussi une dimension qualitative qui permet de mieux comprendre ses effets (King et OPM, 2018, p. 15). Cependant, l'intégration d'une analyse économique traditionnelle dans un devis appliquant l'approche de King est tout à fait possible et peut contribuer à renforcer l'évaluation de la VFM.

Enfin, l'approche de King préconise, avant de poser des jugements sur une intervention, de prendre en compte la capacité d'adaptation de l'organisation qui la met en exécution en analysant l'évolution du contexte dans lequel elle agit ainsi que le niveau de complexité de l'intervention en question. Par ailleurs, l'auteur mentionne l'importance de la prise en compte des différentes perspectives des parties prenantes (donateurs, exécutants, bénéficiaires, etc.), qui peuvent influencer l'élaboration du devis d'évaluation, notamment la sélection des critères, les standards et la méthodologie (King et OPM, 2018, p. 19). L'approche tend à privilégier la perspective des bénéficiaires de l'intervention. Finalement, elle favorise une intégration du suivi et de l'évaluation et, par conséquent, préconise que les cadres soient harmonisés et qu'il y ait une coordination dans la collecte de l'information. Le schéma qui suit résume le cadre conceptuel de l'approche.





**Figure 1**

Schéma conceptuel de l'approche de la VFM de King.

## Une mise en œuvre de l'approche en huit étapes

Pour élaborer un devis et implanter un cadre pour déterminer la VFM, l'application de l'approche de King, en collaboration avec *Oxford Policy Management* (OPM) pour une évaluation robuste, implique une exécution en huit étapes. Les quatre premières concernent l'élaboration d'un devis évaluatif de la VFM et les quatre dernières concernent le rapport évaluatif (King et OPM, 2018, p. 22-23).

- **Étape 1 : Développer la théorie de changement.** Cette étape, effectuée préalablement à la détermination des critères de mérite, des standards de performance et des indicateurs, permet de s'assurer que ceux-ci refléteront la théorie du programme, renforçant ainsi la cohérence de l'évaluation.
- **Étapes 2 et 3 : Développer et définir les critères de jugement et les standards de performance.** Pour réaliser ces deux étapes, il faut, dans l'ordre, rédiger une définition spécifique et concise des cinq critères de mérite, identifier des sous-critères reflétant des dimensions d'une bonne performance pour chacun d'entre eux, et organiser ces sous-critères en niveaux (voir tableau 1 ci-dessus).

- **Étape 4 : Déterminer les indicateurs et les outils de collecte de données.** C'est à cette étape que la question de la mesure se pose pour la première fois. Celle-ci a pour objectif de déterminer les données qui devront être collectées afin d'être en mesure de disposer d'indicateurs qualitatifs et quantitatifs qui reflètent les dimensions des sous-critères sélectionnés préalablement, ce qui permettra par la suite de poser des jugements sur les standards de performance et les critères.
- **Étape 5 : Collecter les données.** Cette étape peut être facilitée si la collecte des données de l'évaluation est alignée avec des données existantes de suivi.
- **Étapes 6 et 7 : Analyser, synthétiser et poser des jugements.** À partir des données collectées, il faut, selon l'ordre suivant, analyser de manière séparée les données provenant des méthodes qualitatives et quantitatives, les synthétiser en triangulant les résultats obtenus des deux analyses et finalement, poser des jugements sur les standards de performance et les critères.
- **Étape 8 : Terminer la rédaction du rapport évaluatif.** Cette étape, une fois effectuée, permet une transmission de l'ensemble des résultats de l'analyse et des constats qui ont été établis par rapport à la VFM de l'intervention (voir tableau 2 ci-dessus).

Le cadre dans lequel s'inscrivent ces étapes va au-delà de la prise en compte d'une question évaluative du type : « Dans quelle mesure l'intervention est-elle efficiente et comment cette efficacité peut-elle être améliorée? », pour répondre à la question plus large suivante : « Quelle est la valeur de l'intervention au regard des critères suivants précisément définis pour un contexte donné : économique (rapport qualité-prix); efficacité; efficacité (en termes de résultats); équité (justice sociale); coût-efficacité (quel est l'impact de l'intervention obtenu avec le niveau d'investissement financier engagé)? »

### Cas d'application du cadre de King : le MUVA, un programme d'autonomisation économique des femmes au Mozambique

Julian King a déjà appliqué lui-même son cadre dans le contexte de différents programmes, dont le MUVA, un programme financé par le Département britannique pour le développement international (DFID), dirigé par l'*Oxford Policy Management*, et qui finance différents projets novateurs pouvant contribuer à accroître l'autonomie économique des jeunes femmes vivant en milieu urbain au Mozambique (King et OPM, 2018, p. 5). Évaluer la VFM pour ce programme constituait un défi important étant donné la nature du programme et la difficulté à donner une valeur monétaire à ses effets intangibles tels que la qualité de vie, l'autonomie et la confiance en soi. King et Guimaraes (2016) ont donc suivi les étapes de l'approche, notamment en établissant des critères de jugement et des standards de performance spécifiques et adaptés au contexte de MUVA en lien avec les « 4 E ». Par exemple, en ce qui concerne l'efficacité, trois critères qui découlent de la théorie de changement ont été définis, soit l'efficacité en tant que programme d'autonomisation économique des femmes, d'apprentissage et d'influence auprès des différentes parties prenantes (King et OPM, 2018, p. 63). À chacun de ces critères, des sous-critères reflétant les différentes dimensions de performances ont été élaborés. Notamment, en lien avec son efficacité en tant que programme d'influence, les sous-critères qui ont été utilisés sont les suivants : son influence auprès des autres programmes du DFID, le niveau d'implantation de ses approches les plus performantes par ses partenaires et ses impacts sur les parties prenantes qui sont censées devenir d'importants agents de changement (King, 2019b, p. 220). Pour chacun d'entre eux, des standards de performance et des rubriques ont été développés en collaboration avec les différentes parties prenantes. Par la suite, différents indicateurs qualitatifs et quantitatifs y ont été



intégrés pour chacun des critères. Par exemple, pour l'efficacité en tant que programme d'apprentissage, les auteurs ont utilisé deux indicateurs qualitatifs en lien avec la perception des parties prenantes par rapport à la participation locale, sur les connaissances acquises et les apprentissages découlant des différents projets, ainsi qu'un indicateur quantitatif visant le nombre de rapports issus des rencontres réalisées et qui mentionne des impacts positifs des projets (King et Guimaraes, 2016, p. 63). En bref, l'utilisation du cadre a permis de fournir des éléments de réponses plus précis et des explications en lien avec les différents concepts d'économie, d'efficience, d'efficacité et d'équité qui n'auraient pu être approfondis dans le cadre d'une analyse économique.

## Conclusion

L'approche VFM de King revisite les enjeux de l'optimisation des ressources investies dans les interventions de développement ou encore les questions du retour sur investissement (ROI en anglais). Si les approches économiques en évaluation y ont apporté certaines réponses, il demeure que la complexité de ces interventions, amplifiée par le manque de données notamment statistiques fiables, par l'absence d'une pratique de documentation de l'historique d'investissement des acteurs, par la difficulté de trouver des comparatifs dans des contextes aussi diversifiés, etc., limitent leur portée et leur utilisation. En proposant un cadre d'évaluation plus élargi pour déterminer la « valeur » d'une intervention au regard de critères tels que le rapport qualité-prix, l'efficience ou encore l'équité, King et l'OPM (2018), s'inscrivant dans la réflexion de Jackson (2012a, 2012b), nous amènent à nous poser les questions suivantes : de quelle intervention ou composante d'intervention s'agit-il ? De quelle valeur parle-t-on ? Selon la perspective de qui ? Et sur quelle période ? Évaluer la performance d'une intervention intègre plusieurs dimensions et exige de faire preuve de pensée évaluative dans la détermination de celle-ci.

## Pour aller plus loin

- King, J. (2017). Using economic methods evaluatively. *American Journal of Evaluation*, 38(1), 101-113. Cet article soutient que lorsque l'on aborde une question évaluative d'un point de vue économique (le mérite, la valeur ou l'importance de l'utilisation des ressources), cette perspective peut améliorer l'évaluation, mais est généralement insuffisante pour répondre pleinement à la question d'évaluation. Pour répondre à cet enjeu, l'auteur propose une approche plus intégrée impliquant un raisonnement évaluatif plus robuste appuyé par une utilisation judicieuse de méthodes mixtes incluant les méthodes économiques. Un cadre théorique global est proposé et les implications pour la pratique de l'évaluation sont traitées.
- King, J. et Allan, S. (2018). Applying evaluative thinking to value for money : The Pakistan Sub-national Governance Programme. *Evaluation Matters—He Take Tō Te Aromatawai*, 4, 207-235. Dans cette étude de cas d'un programme de gouvernance infranationale au Pakistan (organisme de soutien aux réformes sur la gestion des finances publiques, à la planification et à la réalisation de projets-pilotes innovants visant l'amélioration des services), King fait la démonstration que son approche est pertinente pour relever les défis qui se posent lorsque l'on évalue la VFM dans des programmes complexes et adaptatifs et que les principes de base de la démarche sont transférables à d'autres secteurs.

- Jackson, P. (2012b). *Value for money and international development : Deconstructing myths to promote a more constructive discussion*. IDEV. Accessible ici : [http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Evaluation%20Matters%20Q3\\_EN.pdf#page=14](http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Evaluation%20Matters%20Q3_EN.pdf#page=14). Ce rapport propose de dissiper la confusion qui peut être faite autour de la notion d'optimisation des ressources (VFM en anglais) et plaide en faveur d'une discussion plus constructive de sa pertinence et de ses limites en contexte de développement.
- Adou, J.-Y. (2016). Applications et limitations de l'optimisation des ressources dans les activités de développement. *eVALUation Matters/Troisième trimestre*. Accessible ici : <http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Applications%20et%20limitations%20de%20loptimisation%20des%20ressources.pdf>. L'auteur propose ici une réflexion autour de la problématique de l'allocation optimale des ressources d'aide ainsi que de l'efficacité de l'aide au développement.

## Références bibliographiques

- Adou, J.-Y. (2016). Applications et limitations de l'optimisation des ressources dans les activités de développement. *eVALUation Matters/Troisième trimestre*. <http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Applications%20et%20limitations%20de%20loptimisation%20des%20ressources.pdf>
- Brousselle, A., Lachaine, J. et Contandriopoulos, A.-P. (2018). L'évaluation économique. Dans F. Champagne et Z. Hartz (dir.), *L'évaluation : concepts et méthodes* (p. 199-235). Presses de l'Université de Montréal. <http://books.openedition.org/pum/6310>
- Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J. et Stoddard, G. L. (2005). *Methods for the economic evaluation of health care programs*. Oxford University Press.
- Fleming, F. (2013). *Evaluation methods for assessing Value for Money*. BetterEvaluation. <https://www.betterevaluation.org/sites/default/files/Evaluating%20methods%20for%20assessing%20VfM%20-%20Farida%20Fleming.pdf>
- Fox-Rushby, J. et Cairns, J. (2005). *Economic evaluation*. McGraw-Hill Education.
- Glendinning, R. (1988). The concept of value for money. *International Journal of Public Sector Management*, 1(1) 42-50.
- IDEV. (2016). Value for money in development work. *eVALUation Matters, Third Quarter 2016*. [http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Evaluation%20Matters%20Q3\\_EN.pdf](http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Evaluation%20Matters%20Q3_EN.pdf)
- Jackson, P. (2012a). Value for money and international development : Le Comité d'aide au développement (CAD) de l'OCDE.
- Jackson, P. (2012b). *Value for money and international development : Deconstructing myths to promote a more constructive discussion*. IDEV. [http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Evaluation%20Matters%20Q3\\_EN.pdf#page=14](http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Evaluation%20Matters%20Q3_EN.pdf#page=14)
- King, J. (2017). Using economic methods evaluatively. *American Journal of Evaluation*, 38(1), 101-113.

- King, J. (2019a). *Evaluation and Value for Money : Development of an approach using explicit evaluative reasoning* [thèse de doctorat, Université de Melbourne]. [https://minerva-access.unimelb.edu.au/bitstream/handle/11343/225766/PhD %20thesis %20Julian %20King %20July %202019.pdf? sequence=3&isAllowed=y](https://minerva-access.unimelb.edu.au/bitstream/handle/11343/225766/PhD%20thesis%20Julian%20King%20July%202019.pdf?sequence=3&isAllowed=y)
- King, J. (2019b). Combining multiple approaches to valuing in the MUVA female economic empowerment programme. *Evaluation Journal of Australasia*, 19, 217-225. <https://doi.org/10.1177/1035719X19893080>
- King, J. (2020, février). *Evaluation and Value for Money : A practical approach*. Australasian Aid Conference, Canberra.
- King, J. et Allan, S. (2018). Applying evaluative thinking to value for money : The Pakistan Sub-national Governance Programme. *Evaluation Matters—He Take Tō Te Aromatawai*, 4, 207-235.
- King, J. et Guimaraes, L. (2016). Evaluating value for money in international development : the Ligada female economic empowerment programme in Mozambique. *EVALUation Matters, Third quarter*. [http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Evaluating %20 value %20for %20money %20in %20international %20development-.pdf](http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Evaluating%20value%20for%20money%20in%20international%20development-.pdf)
- King, J., McKegg, K., Oakden, J. et Wehipeihana, N. (2013). Evaluative rubrics : A method for surfacing values and improving the credibility of evaluation. *Journal of MultiDisciplinary Evaluation*, 9(21), 11-20.
- King, J. et OPM VfM Working Group. (2018). *OPM's approach to assessing Value for Money*. *Oxford Policy Management*. <https://www.julianking.co.nz/wp-content/uploads/2018/02/OPM-approach-to-assessing-value-for-money.pdf>



# La modélisation des interventions en évaluation



*Juan Manuel Lopez Mejia*  
*Lynda Rey*

## Introduction

La modélisation des interventions est l'une des étapes fondamentales de la démarche évaluative. Avant de procéder à l'évaluation d'une intervention, il est important de bien comprendre les contours et les limites de celle-ci. La modélisation est une représentation de la théorie du changement<sup>1</sup> sous-jacente à l'intervention (Abdi et Mensah, 2016). Elle permet de résumer les éléments essentiels d'une intervention à travers une relation « si...alors » faisant gage de la logique qui associe les choix d'activités et les résultats qui doivent en découler. Elle repose sur l'hypothèse suivante : *si* les enchaînements logiques de l'intervention sont appliqués comme prévu, *alors* les changements souhaités ont plus de chances de se concrétiser. Plusieurs outils sont utilisés pour modéliser l'intervention, mais les plus couramment employés sont le modèle logique et la théorie de changement (ToC) (Mayne, 2017). Ceux-ci sont souvent cités de façon interchangeable dans la littérature scientifique et professionnelle, mais il existe quelques nuances qui permettent de les différencier. Tandis que le premier vise à mettre en évidence les liens logiques des différents éléments généralement de façon linéaire, le second va plus loin en intégrant à ces liens logiques les hypothèses, les risques ainsi que les facteurs internes et externes à l'intervention. Malgré ces différences, le modèle logique et la théorie de changement partagent les mêmes utilités et accompagnent le cycle de l'intervention, de la planification à l'évaluation en passant par la mise en oeuvre. Ils constituent également d'excellents outils de communication. En somme, ils font partie de la boîte à outils des spécialistes de l'évaluation. La personne évaluatrice doit les maîtriser pour s'assurer de la qualité de sa démarche d'évaluation.

Dans un premier temps, ce chapitre définit le modèle logique. Il explique son utilité pour la planification, l'évaluation et la communication, son processus de construction et de présentation ainsi que la question du contrôle et de l'influence. Dans un deuxième temps, il décrit la théorie de changement en s'appuyant sur sa relation avec le modèle logique. Cette partie définit les hypothèses et les risques ainsi que les facteurs internes et externes, tout en mettant de l'avant les avantages qui découlent de l'élaboration de la théorie de changement d'une intervention. À la fin de ce chapitre, un tableau de validation de la qualité d'un modèle logique et d'une ToC sera proposé.

---

1. Dans ce chapitre, le terme « changement » est utilisé de façon générique et peut signifier les résultats/effets d'une intervention.

## Qu'est-ce qu'un modèle logique ?

Le modèle logique consiste dans la représentation schématique des éléments constitutifs d'une intervention : ressources ou intrants, activités, produits ou extrants, groupes cibles, et ses résultats visés à court, moyen et long terme (Tsala, 2019; W.K. Kellogg Foundation, 2004). Une intervention peut être construite de façon restreinte ou élargie, mais elle est fondamentalement composée d'une série d'activités, elles-mêmes appuyées sur un nombre de ressources, qui ont pour but l'atteinte d'un ou de plusieurs résultats (Porteous, 2013). Il faut noter ici que les résultats renvoient aux effets attendus de la part des personnes responsables de la conception des interventions. Cet outil vise avant tout à mettre en évidence les liens logiques du type « si... alors » existant entre ces différents éléments. Pour que ces relations puissent être démontrées, il est indispensable que l'intervention à mettre en place s'appuie sur une *théorie de programme*. Cette dernière repose sur le « raisonnement explicite [quant au] comment le changement aura lieu et comment l'intervention est censée produire les liens causaux » (Funnell et Rogers, 2011, p. 13; Rogers, 2016). Pour trouver une théorie adaptée à l'intervention que l'on souhaite implanter, encore faut-il réfléchir sur les différentes théories existantes les plus pertinentes pour refléter les intentions des concepteurs (théories du développement institutionnel, théorie de la mobilisation communautaire, théorie du changement individuel, etc.). Sans la théorie de programme, il est difficile de démontrer que l'intervention atteindra les résultats escomptés<sup>2</sup> (Rogers, 2016).

Par ailleurs, pour certains auteurs tels que Chen *et al.* (2018), Pawson et Tilley (1997), Funnell et Rogers (2011), Brousselle et Buregeya (2018), la théorie de programme se compose de la *théorie de changement* et de la *théorie de l'action* (Montague, 2019). Tandis que la première explique *pourquoi* l'intervention est censée fonctionner, la seconde s'intéresse à *comment* celle-ci doit être mise en œuvre. Les questions liées à la conception, à la planification, et à la mise en œuvre font partie de la théorie de l'action. Toute intervention se déploie dans un contexte donné qui sous-entend l'étroite relation de l'intervention avec son contexte (voir les approches d'évaluation basées sur la théorie). Ce serait donc la combinaison des caractéristiques contextuelles, des processus de mise en œuvre et des relations causales qui reflètent comment *les produits des activités* pourraient mener à des changements à court, moyen et long terme.

### Du cadre logique au modèle logique

Le modèle logique est le résultat de l'évolution des pratiques évaluatives. Ses origines remontent en 1969 lorsque l'Agence américaine pour le développement international (*USAID*) a adopté l'outil-cadre logique (ou *LogFrame* en anglais) pour l'élaboration, la planification, le suivi et l'évaluation des programmes de développement (Dhillon et Vaca, 2018). Dans sa version la plus simple, il s'agit d'une matrice représentant à la verticale la logique d'intervention. Cette dernière est composée de quatre ou cinq lignes, les activités, les produits et les objectifs spécifiques et globaux à atteindre. À l'horizontale, la matrice est composée de trois ou quatre colonnes représentant les indicateurs objectivement vérifiables, les sources de vérification, les hypothèses et les risques (voir figure 1). Jusque dans les années 90, le cadre logique a été largement utilisé dans le secteur du développement international, aussi bien par la Banque mondiale que par les organisations non gouvernementales (ONG), comme un prérequis pour l'attribution de financement (Ringhofer et Kohlweg, 2019).

2. C'est ce que plusieurs auteurs appellent le problème de la boîte noire (*black box*), qui se caractérise par une absence d'information sur le déroulement du programme (Funnell et Rogers, 2011).

Logique d'intervention	Indicateurs objectivement vérifiables	Sources de vérification	Hypothèses et risques <sup>3</sup>
<b>Objectif global</b> Réduire de 50 % les discriminations liées au genre au sein des formations sanitaires en milieu urbain et en zone rurale en 5 ans.	Pourcentage des discriminations liées au genre dans les centres de santé en milieu urbain et en zone rurale.	Registre des plaintes	S. O.
<b>Objectif spécifique</b> Améliorer les compétences du personnel de santé en matière de prise en compte des approches liées au genre dans sa pratique quotidienne.	Compétences liées au genre du personnel de santé en milieu urbain et en zone rurale.	Test de compétences postformation Rapport de supervision	L'amélioration des compétences renforce la confiance du personnel de santé et favorise l'accès équitable aux soins de santé.
<b>Réalisations</b> 200 membres du personnel infirmier et médical en milieu rural et 500 membres en milieu urbain ont participé à 2 formations de 5 jours par année.	Nombre de personnes ayant complété les formations. Nombre de formations offertes annuellement.	Rapport de formation	Le personnel formé a acquis de nouvelles connaissances et compétences pour appliquer les approches liées au genre.
<b>Activités</b> Former le personnel de santé aux approches liées au genre dans les centres de santé.	Nombre de personnes formées dans les centres de santé.	Rapport d'activités	Le personnel de santé est disponible pour participer à la formation.

**Figure 1**

### Structure du cadre logique

Source : d'après un cas issu de l'expérience des auteurs.

La rigidité de cet outil et l'absence de relations causales entre les éléments de la matrice du cadre logique ont conduit à l'adoption d'une version simplifiée du cadre logique que l'on nomme désormais modèle logique (*Logic Model*) par l'organisme *United Way of America* en 1996 (Chen *et al.*, 2018), qui s'est dès lors imposé comme un outil indispensable pour l'évaluation. L'évolution de la représentation des modèles logiques se poursuit de nos jours. En effet, selon le type d'intervention, les spécialistes en évaluation intègrent de plus en plus dans leurs modèles des éléments transversaux tels que le genre, les droits de la personne, le développement durable, etc.

3. À noter que les risques et hypothèses peuvent également être évalués afin de mieux comprendre les répercussions d'une intervention.



## Un outil de planification, de communication et d'évaluation

La création d'un modèle logique comporte de nombreux avantages, notamment pour la planification, la gestion, la communication et l'évaluation d'une intervention. L'utilisation du modèle logique n'est pas exclusive à la phase évaluative d'une intervention. En effet, cet outil visuel est devenu indispensable dans la phase de planification, notamment parce qu'il permet de comprendre les tenants et aboutissants d'une intervention. L'outil précise les besoins justifiant son existence et sa raison d'être. Tout travail d'évaluation d'une intervention peut démarrer par la (re)construction (analyse de l'évaluabilité<sup>4</sup>), à la validation ou encore à la mise à jour du modèle logique. Dès lors, certains modèles incluent la mission de l'organisme et la raison d'être d'une intervention<sup>5</sup>.

Le modèle logique est également un outil primordial pour la communication. Telle une feuille de route, il permet de communiquer une vision commune sur les besoins ainsi que sur le cheminement d'une intervention; des activités aux résultats à atteindre. Il permet aux parties prenantes<sup>6</sup> et aux partenaires d'observer en un coup d'œil la raison d'être ainsi que l'enchaînement logique des étapes de façon synthétique et transparente. Le modèle logique réduit les interprétations sur l'intervention, et rend explicite la théorie sous-jacente de l'intervention, notamment lorsque cette dernière comporte un degré de complexité (Rey *et al.*, 2012). Bien que le modèle logique constitue une simplification de la réalité, il n'a pas pour intention de minimiser la complexité de l'intervention. Selon Patton (2011), « face à la complexité, la première tâche [des spécialistes de l'évaluation et de ses partenaires] est d'identifier des objectifs clairs, spécifiques et mesurables... tout semble complexe jusqu'à ce que l'on construise un modèle logique » (Patton, 2011, p. 6, cité dans Rey *et al.*, 2012, p. 81; librement traduit par les auteurs).

Rappelons que l'existence d'un modèle logique ou d'une théorie de changement (ToC) témoigne « d'une vision explicite (et donc évaluable) de la façon *dont on pense* que le changement induit par un projet ou programme doit se produire » (Davis, 2008, cité dans Tsala, 2019). Dès lors que le modèle logique d'une intervention est disponible, les spécialistes de l'évaluation ont entre leurs mains un outil précieux leur permettant de vérifier si les changements ou effets attendus se sont effectivement produits. Si tel n'est pas le cas, les parties prenantes, en collaboration avec les spécialistes en évaluation, pourront alors réfléchir aux causes de non-survenue des effets, et seront à même d'ajuster l'intervention pour que les résultats soient atteints. Dans le tableau suivant, nous énumérons un certain nombre d'avantages du modèle logique :

4. L'analyse de l'évaluabilité consiste en la reconstruction de l'histoire de l'intervention. Cette reconstruction permet le rétablissement de la logique de cette dernière et facilite la définition de l'objet à évaluer. «Après tout, la discipline de l'évaluation vise essentiellement à déterminer si cette logique liant les activités d'une intervention à ses résultats anticipés s'est matérialisée comme prévu » (Dumaine, 2012, p. 98). À noter également que l'analyse de l'évaluabilité peut également être une évaluation de l'état de préparation d'un programme à recevoir une évaluation.
5. Voir le modèle logique élaboré par le partenariat *Filles, pas épouses* (2020). <https://www.girlsnotbrides.org/documents/1311/Theorie-du-Changement-sur-le-Mariage-des-Enfants-Brochure-1.pdf>
6. Dans ce chapitre, les parties prenantes incluent les bénéficiaires de l'intervention.

**Tableau 1 — Avantages d'un modèle logique. Inspiré de Porteous, 2013 et Dillion et Vaca, 2018.**

Phases du programme	Avantages
<b>Planification</b>	<ul style="list-style-type: none"> <li>▪ Faciliter l'établissement d'une mission commune;</li> <li>▪ Réduire les interprétations sur les résultats à atteindre;</li> <li>▪ Aligner l'intervention avec la mission de l'organisme;</li> <li>▪ Faciliter l'engagement de l'équipe à travers sa participation dans la planification;</li> <li>▪ Relier la planification stratégique à la planification opérationnelle;</li> <li>▪ Faciliter l'établissement d'indicateurs de rendement/de performance;</li> <li>▪ Faciliter la demande de subventions auprès des bailleurs de fonds;</li> <li>▪ Communiquer avec les parties prenantes et les groupes cibles les raisons d'être de l'intervention;</li> <li>▪ Faciliter l'évaluation de l'intervention.</li> </ul>
<b>Mise en œuvre</b>	<ul style="list-style-type: none"> <li>▪ Faciliter le suivi (<i>monitoring</i>) des activités;</li> <li>▪ Éclairer la prise de décision pendant la mise en œuvre;</li> <li>▪ Rappeler les intentions et la raison d'être de l'intervention;</li> <li>▪ Permettre d'ajuster les activités et extrants lorsque des situations complexes émergent.</li> </ul>
<b>Évaluation</b>	<ul style="list-style-type: none"> <li>▪ Ajuster les stratégies de façon à atteindre les résultats escomptés;</li> <li>▪ Rendre l'intervention compréhensible auprès du public cible;</li> <li>▪ Associer les résultats en fonction des activités et extrants de l'intervention;</li> <li>▪ Comprendre les raisons pour lesquelles les effets de l'intervention ne se sont pas produits;</li> <li>▪ Permettre d'identifier les succès et défis d'une intervention et leur contribution au changement;</li> <li>▪ Démystifier l'évaluation en facilitant la participation et inciter à l'utilisation des résultats de l'évaluation.</li> </ul>

### Le modèle logique comme processus participatif

Avant d'aborder les différentes parties du modèle logique, il est important de souligner que la construction de celui-ci doit être participative. L'inclusion des parties prenantes (partenaires, bénéficiaires, employés) est garante d'une vision commune et partagée qui guidera l'intervention vers l'obtention des résultats. La personne chargée de réaliser l'évaluation doit reconnaître que ses propres perspectives et croyances déterminent sa façon de comprendre comment le changement est en train de se produire. De même, sa propre perception du monde influence la façon dont elle décide de capturer et d'analyser ce changement (Dhillon et Vaca, 2018). Elle doit donc être consciente de ses propres biais cognitifs, autrement dit, l'ensemble des croyances et des valeurs qui peuvent influencer son jugement (Chaumont et Leroux, 2018).

Traditionnellement, les spécialistes de l'évaluation développaient le modèle logique de façon indépendante à l'aide de quelques ressources documentaires et d'entretiens réalisés auprès du personnel d'un organisme. Aujourd'hui, il est fortement recommandé de travailler dès le départ avec les parties prenantes de l'intervention dans l'élaboration du modèle logique. Cet exercice est gage de transparence

de la démarche d'évaluation. Par la même occasion, il réduit les diverses interprétations sur la stratégie à entreprendre, favorise l'adoption d'une vision commune, contribue au renforcement des capacités évaluatives des parties prenantes et facilite l'évaluation. L'apport de nouvelles connaissances sur l'évaluation permet aux parties prenantes de comprendre et d'appréhender différemment certains aspects de l'intervention (utilisation conceptuelle). De même, la participation à l'élaboration du modèle logique influencera les façons de penser et de faire des parties prenantes, et pourra ultimement favoriser la prise de décision quant à la structure ou à l'orientation de l'intervention (utilisation processuelle) (Alkin et King, 2016; Alkin et Taut, 2003; Forss *et al.*, 2002).

Au moment de l'élaboration de cet outil, il incombe aux spécialistes de l'évaluation d'utiliser leurs capacités de facilitation/d'animation pour inciter les personnes impliquées à « vérifier l'exactitude des travaux; reconnaître la nécessité de contester ou de remettre en question la logique [de l'intervention]; et obtenir un accord sur le modèle logique établi avant de procéder à l'évaluation » (Porteous, 2013, p. 102). La démarche participative peut être longue et coûteuse, mais elle suscite l'engagement et l'appropriation des résultats par les parties prenantes.

Rappelons enfin que l'élaboration d'un modèle logique demande plusieurs essais et itérations. D'après nos expériences, il est important que la personne ou l'équipe chargée de l'évaluation réalise un travail préliminaire en amont (examen des documents et des données administratives, entretiens préliminaires sur le programme, examen de la littérature) avant d'entreprendre un processus d'élaboration d'un modèle logique. Par ailleurs, il est rare que l'on aboutisse à une version finale en une seule tentative. Souvent, il faut plusieurs rencontres avec les parties prenantes pendant lesquelles l'ébauche sera retravaillée. Il est donc très important de prendre le temps d'aborder tous les aspects du programme, de discuter de toutes les perspectives des participants et d'obtenir le consensus sur une version définitive.

### Construire un modèle logique : les éléments constitutifs

Les éléments d'un modèle logique peuvent varier d'un organisme à l'autre. Toutefois, nous considérons qu'un bon modèle doit représenter au minimum les éléments suivants :

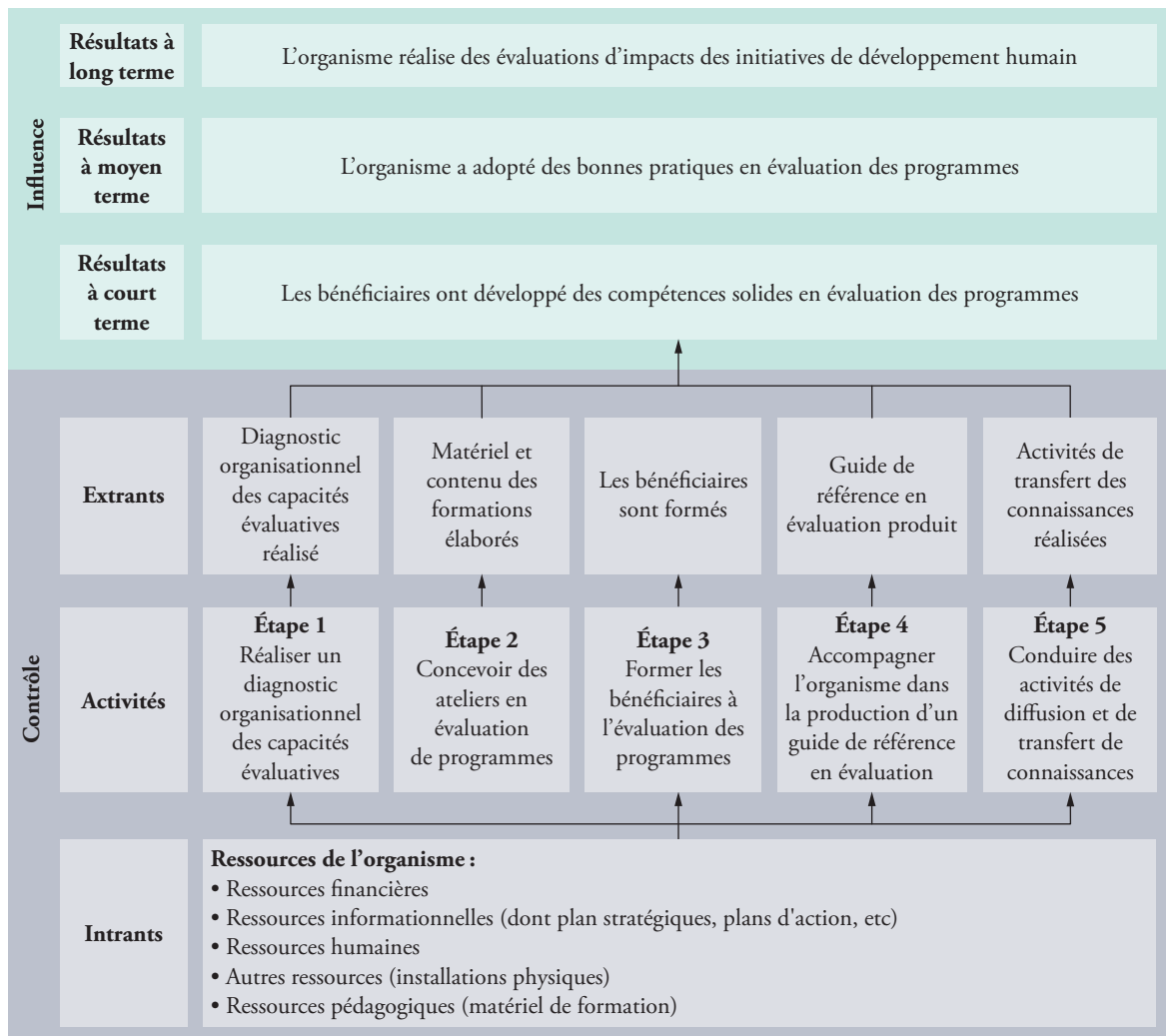
- a. Les **intrants (ou ressources)** sont toutes les ressources nécessaires à la réalisation des activités. Dans cette rubrique sont incluses les ressources financières (budget, fonds), les ressources humaines (personnel, partenariats, experts, bénévoles, etc.), les ressources informationnelles (planification stratégique, documents de formation, les ententes, etc.) et les ressources matérielles et physiques (équipements, locaux, etc.). En somme, les intrants constituent toutes les ressources qui permettront de mettre en œuvre les activités visant à l'atteinte des résultats de l'intervention.
- b. Les **activités** constituent toutes les actions entreprises par l'organisme pour transformer les ressources en produits. Ce sont les fonctions que l'organisation accomplit dans son quotidien (former le personnel, construire une école, sensibiliser une population cible). Les activités sont toujours formulées en verbes d'action (former, distribuer, construire, sensibiliser).
- c. Les **populations (ou groupes) cibles** sont tous les individus ou organismes bénéficiaires (directs et indirects) de l'intervention. Elles peuvent être déterminées en fonction des caractéristiques sociodémographiques ou en fonction des problématiques auxquelles le programme souhaite s'adresser (Porteous, 2013). En fin de compte, c'est au sein de cette population cible que l'on voudrait observer des changements des conditions sociales ou économiques. Il peut y avoir plusieurs groupes cibles selon l'intervention concernée.

- d. Les **extrants (ou produits)** représentent les biens ou les services issus d’une production. En d’autres termes, ce sont les résultats opérationnels des actions entreprises par l’organisme auprès des bénéficiaires de l’intervention. En guise d’exemple, si notre activité visait à « former le personnel d’un organisme partenaire », notre extrant est alors « personnel de l’organisme partenaire formé ». Les extrants sont concrets et peuvent être quantifiés.
- e. Les **résultats** sont les changements que l’intervention souhaite créer. Ils constituent la raison de l’existence même de l’intervention mise en place. Ces changements peuvent survenir au niveau des capacités, des connaissances, des aptitudes, des comportements et des changements d’état, par exemple (voir tableau 2 ci-dessous). Les résultats sont généralement divisés en **résultats à court, moyen et long terme**. Ce découpage des résultats permet de mieux apprécier les différents niveaux de changement.

Notons également que pour l’élaboration du modèle logique, il est préférable d’adopter une approche rétrospective. Cela signifie que l’on identifie d’abord les populations cibles, on définit ensuite les résultats à long, moyen et à court terme, les produits, les activités et les ressources. L’avantage de l’approche rétrospective est qu’elle permet de visualiser les résultats que l’on souhaite atteindre et de construire l’intervention en fonction de ces derniers (Community Sustainability Engagement Evaluation Toolbox, 2010).

**Tableau 2 — Les changements escomptés par niveau des résultats. Inspiré de Porteous, 2013.**

Résultats	Changements
<b>Résultats à court terme</b>	Changements relatifs au degré de sensibilisation à une problématique, aux connaissances, aux aptitudes ou aux compétences. Changements du point de vue de : <ul style="list-style-type: none"> <li>▪ Conscience;</li> <li>▪ Connaissance;</li> <li>▪ Attitudes;</li> <li>▪ Compétences;</li> <li>▪ Opinion;</li> <li>▪ Motivation;</li> <li>▪ Aspiration;</li> <li>▪ Motif;</li> <li>▪ Intention.</li> </ul>
<b>Résultats à moyen terme</b>	Changements d’habitudes ou de comportements. Changements du point de vue de : <ul style="list-style-type: none"> <li>▪ Pratiques;</li> <li>▪ Actions sociales;</li> <li>▪ Décisions politiques.</li> </ul>
<b>Résultats à long terme</b>	Changements des conditions sociales, économiques, environnementales ou liées à la santé. Changements du point de vue de : <ul style="list-style-type: none"> <li>▪ Conditions sociales (bien-être);</li> <li>▪ Sanitaires;</li> <li>▪ Économiques;</li> <li>▪ Environnementales.</li> </ul>



**Figure 2**

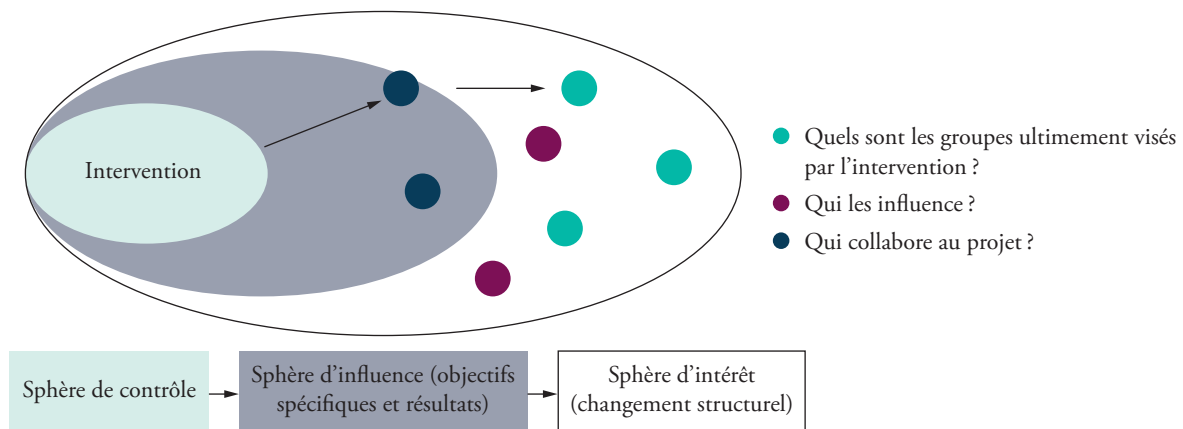
Modèle logique d'une intervention en renforcement des capacités évaluatives<sup>7</sup>.

Source : Lopez, Messomo Beyene et El Kaddouri (2020).

7. Ce modèle logique s'est inspiré d'un travail réalisé dans le cadre d'un cours portant sur une démarche de renforcement des capacités organisationnelles en évaluation d'un partenaire par l'ENAP.

## Contrôle et influence

Il est important de faire la différence entre la zone de contrôle et celle de l'influence. En effet, le gestionnaire de l'intervention exerce un contrôle sur les premiers éléments du modèle logique, à savoir les *intrants*, les *activités* et les *extrants*. Il est responsable de son budget et du suivi d'un calendrier d'activités qui mèneront à la production d'extrants. Le gestionnaire et son équipe pourraient, certes, rencontrer certains obstacles pendant la préparation ou la conduite des activités, mais ils ont la responsabilité de « contrôler » les imprévus opérationnels qui peuvent apparaître pendant la mise en œuvre de l'intervention (Porteous, 2013). Après la production des extrants, le gestionnaire et son équipe peuvent, au mieux, espérer « influencer » les *résultats à court, moyen et long terme* (voir figure 2). Cela s'explique par le fait que les bénéficiaires de l'intervention sont soumis à des facteurs externes qui vont nécessairement influencer la façon dont ils interagissent avec l'intervention et, de surcroît, avec ses résultats. Bien que l'on ne puisse pas contrôler tous les aspects de l'intervention, on peut reconnaître *l'influence* (ou la contribution) de celle-ci sur les changements des comportements et des conditions de vie.



**Figure 3**

Schéma des sphères de changement.

Source : adapté de Montague et Birch-Jones, 2017, cité dans Kinyuira, 2019, p. 30.

## Présentation visuelle d'un modèle logique

Notons qu'il n'y a pas une façon standard de présenter un modèle logique. Il existe un grand nombre de styles et de formats. Le point commun des modèles existants réside dans le fait que le modèle logique est généralement présenté sur une page. Chacun des éléments est inséré dans une boîte, souvent de la même taille et dans certains cas, de la même couleur. Les modèles logiques peuvent être illustrés soit de façon verticale (voir figure 3) ou horizontale, selon les préférences des parties prenantes. Les liens causaux entre les différents éléments sont habituellement représentés par des flèches de la même taille (Dhillon et Vaca, 2018).

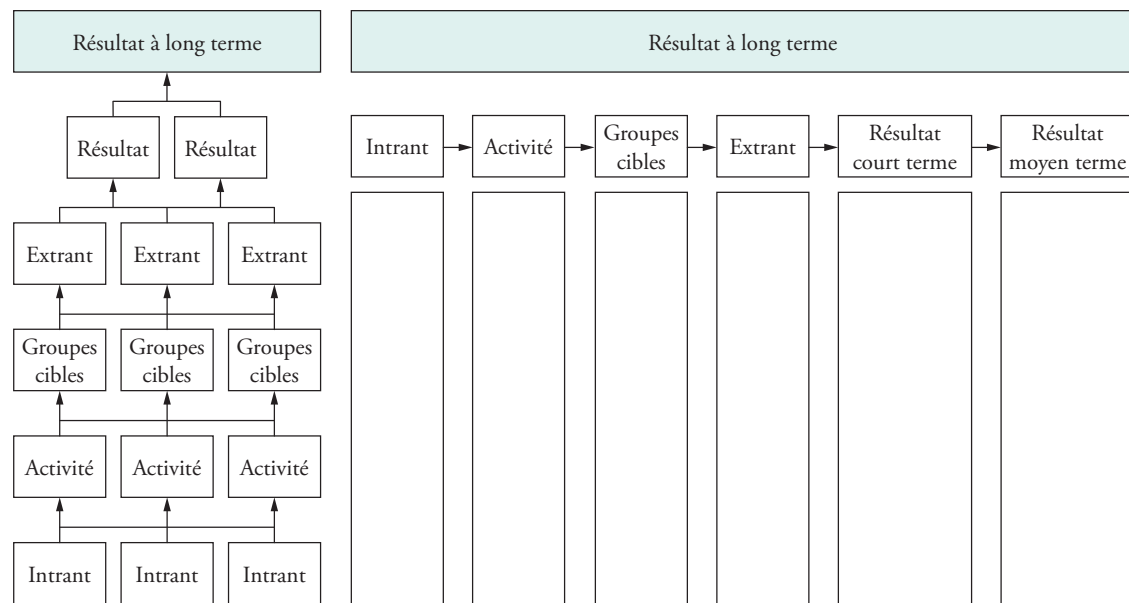


Figure 4

Format des modèles logiques les plus courants

La présentation du modèle logique peut sembler superflue. Néanmoins, comme nous l'avons indiqué plus haut, l'une des raisons d'être du modèle logique est de communiquer les tenants et aboutissants de l'intervention de la façon la plus claire possible. L'ajout des couleurs, des flèches, des boîtes de tailles différentes peuvent s'avérer utiles pour exprimer un degré plus élevé de complexité, que ce soit sur un modèle logique ou une théorie de changement (Dhillon et Vaca, 2018).

## Du modèle logique à la théorie du changement

Les termes théorie du changement (ToC) et modèle logique sont utilisés de façon interchangeable tant dans la littérature scientifique que professionnelle (Dhillon et Vaca, 2018). Cela s'explique par l'absence de consensus sur les éléments caractéristiques de ces deux outils. Les différents écrits placent les origines de la théorie de changement dans la planification des programmes de développement des années 90 (Prinsen et Nijhof, 2015). Elle s'inscrit dans la lignée des approches d'évaluations axées sur la théorie (*Theory-Based evaluation*) et donc, sur le principe selon lequel la contribution d'une intervention aux changements observés repose sur une théorie qui se doit d'être explicite (Secrétariat du Conseil du Trésor du Canada, 2012).

Notons enfin que le modèle logique comporte des limites. Pour certains auteurs, ce dernier reste linéaire et ne permet pas de rendre visible le côté itératif ou complexe de l'intervention (Porteous, 2013). Ce sont des aspects qu'une ToC peut aider à résoudre (Dhillon et Vaca, 2018). La ToC se veut donc un outil plus riche dans le sens où elle vise à compléter le modèle logique par l'ajout de deux éléments supplémentaires, à savoir les hypothèses et les risques (Secrétariat du Conseil du Trésor du Canada, 2012, cité dans TIESS, 2018). Le Conseil du trésor du Québec encourage également l'intégration de la raison d'être et du contexte de l'intervention (Secrétariat du Conseil du trésor du Québec, 2018).

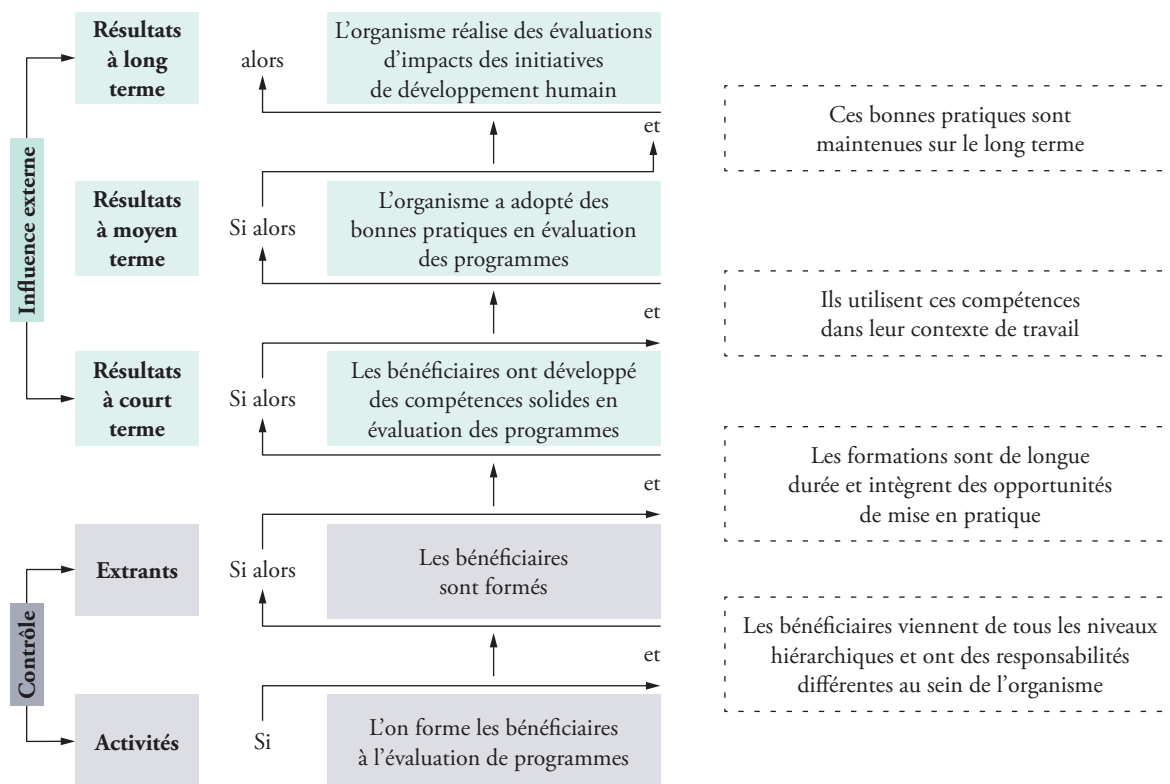
## Les hypothèses

Si le modèle logique désigne les liens logiques entre les différents éléments d'une intervention, la théorie de changement (ToC) désigne, quant à elle, les hypothèses qui expliquent *comment* cette intervention compte atteindre les résultats (Tsala, 2019). La ToC établit les hypothèses causales de chacun des liens. Elles sont généralement représentées par des flèches dans le modèle logique. Le but de préciser les hypothèses consiste à démontrer explicitement *ce qui doit arriver* pour que ces liens de causalité se produisent ou dans *quelles conditions* les changements sont susceptibles de se produire. Comme le suggère Patton (2001), «le fait de préciser les mécanismes de causalité transforme un modèle logique en une théorie du changement» (Mayne, 2017, p. 176).

**Tableau 3** — Les différents types d'hypothèses. Inspiré de Mayne, 2017, p. 124-125.

Hypothèses	Définition
<b>Hypothèses de portée</b>	Conditions nécessaires pour que les extrants découlant des activités soient bien reçus.
<b>Hypothèses de changement de capacité</b>	Conditions nécessaires pour que les changements sur le plan des connaissances se produisent.
<b>Hypothèses de changement de comportement</b>	Conditions nécessaires pour que les changements en matière de comportement aient lieu.
<b>Hypothèses de changement de bien-être</b>	Conditions nécessaires pour que les changements se traduisent par l'amélioration des conditions de vie.





**Figure 5**

### Hypothèses du programme de renforcement des capacités évaluatives

Source : adapté de Mayne, 2017 et Douthwaite, 2020.

Pour construire la figure 4, nous avons repris une activité du modèle logique de la figure 2 et nous avons ajouté les enchaînements causaux et leurs hypothèses. Dans cet exemple, nous avons ajouté une hypothèse par enchaînement logique, mais dans un cas réel, il peut y avoir plusieurs hypothèses pour chaque lien. Il est primordial que l'organisme et la personne chargée de l'évaluation déterminent l'ensemble des hypothèses préalables faisant en sorte que les changements sociaux deviennent une réalité. Les hypothèses sont ici représentées par la conjonction « et » qui permet de compléter l'enchaînement logique de cet élément.

- Ainsi, nous pensons que *si* les bénéficiaires sont formés *et* que ces bénéficiaires viennent des niveaux hiérarchiques supérieurs et ont des responsabilités différentes au sein de l'organisme, *alors* ils développeront des compétences solides en évaluation de programme (hypothèses de portée). L'hypothèse de portée permet également de confirmer les différents groupes et sous-groupes chez les bénéficiaires, et d'analyser les facteurs qui les caractérisent de façon spécifique.
- De même, *si* les bénéficiaires sont formés *et* que les formations sont de longue durée et qu'elles intègrent des opportunités de mise en pratique, *alors* les bénéficiaires développeront des compétences solides (hypothèses de changement de capacité).

- *Si* les bénéficiaires développent des compétences solides *et* qu'ils utilisent ces compétences dans leur contexte de travail, *alors* l'organisation adoptera de bonnes pratiques en matière d'évaluations de programmes (hypothèses de changement de comportement).
- Enfin, *si* l'organisme adopte de bonnes pratiques *et* que celles-ci sont maintenues sur le long terme, *alors* nous pensons que l'organisme réalisera des évaluations d'impacts des initiatives de développement humain<sup>8</sup>.

Comme le décrit Mayne (2017), « l'articulation d'hypothèses sur les liens de causalité implique un savant mélange d'évidences scientifiques, d'expérience des parties prenantes et de théories provenant des sciences sociales » (p. 179). Bien maîtriser la théorie que sous-tend l'intervention est un impératif pour la formulation d'hypothèses solides.

### Les risques

La détermination des hypothèses permet de reconnaître d'éventuels obstacles qui pourraient nuire à l'atteinte des résultats de l'intervention. Chaque hypothèse formulée est en soi un risque pour la réalisation de la théorie de changement (Mayne, 2017). Si nous revenons à notre exemple de la figure 5, nous pourrions conclure que « *si* les formations n'intègrent pas des possibilités de mise en pratique », *alors* les participants ne pourront pas acquérir des connaissances solides. De la même façon, « *si* les participants n'utilisent pas leurs compétences dans leur milieu de travail », *alors* l'organisme n'adoptera pas de bonnes pratiques en matière d'évaluation de programmes. Finalement, « *si* les pratiques ne sont pas maintenues sur le long terme », l'organisme ne sera pas en mesure de réaliser des évaluations efficaces. En définitive, les changements espérés auprès de l'organisme n'auront pas lieu.

### Les effets non souhaités et/ou non prévus

La modélisation des interventions omet souvent de considérer les effets non souhaités et/ou non prévus, positifs ou négatifs (Jabeen, 2018). Or, ceux-ci devraient également être étudiés et suivis dans un contexte donné (Jabeen, 2018; Lemon et Pinet, 2018), d'autant plus que ces effets peuvent, lorsqu'ils sont négatifs, avoir des répercussions néfastes sur les conditions des bénéficiaires au lieu de les améliorer. Selon le modèle de Marceau et collègues, il convient de distinguer autant que possible les effets hors cible, non intentionnels (qui ne correspondent pas aux intentions de l'autorité) et non prévus (qui ne sont pas anticipés par un acteur, actrice).

### Les facteurs externes

Les facteurs externes concernent « les circonstances indépendantes du contrôle du programme », par exemple le contexte sociopolitique ou économique. Ces facteurs viennent influencer directement la réalisation des résultats et le comportement des groupes cibles, et ont donc un effet sur la capacité de l'intervention d'atteindre ses objectifs. Il est important de formuler des hypothèses et de définir les facteurs externes dans le but de placer l'intervention dans son contexte. Cet exercice, effectué à l'étape de conception de l'intervention, permet d'identifier des activités additionnelles qui peuvent s'avérer utiles dans le cadre de la gestion des risques (Secrétariat du Conseil du Trésor du Canada, 2012).

---

8. Étant donné que l'organisme partenaire est la cible de l'intervention, nous ne pouvons pas parler d'une hypothèse de changement de bien-être.

## Les avantages de réaliser une théorie de changement

Comme nous l'avons indiqué pour le modèle logique, la construction d'une ToC s'avère très utile pour la planification, pour la communication des tenants et aboutissants, et bien évidemment pour l'évaluation. Pour les gestionnaires de l'intervention, la ToC leur offre la possibilité de tester des stratégies cohérentes en même temps qu'elle garantit le partage d'une vision commune entre eux et les employés de l'organisme. Pour les employés, cet outil est une feuille de route pour la planification et la mise en œuvre d'interventions spécifiques. Enfin, pour les différentes parties prenantes, la ToC permet de comprendre facilement la raison d'être de l'intervention. Elle indique la manière dont les différents partenaires peuvent mettre à profit leurs ressources et leurs forces à l'atteinte des résultats (Dhillon et Vaca, 2018). Enfin, pour les personnes chargées de l'évaluation, la ToC facilite la compréhension de l'intervention et la mise en place d'évaluations futures.

## La présentation d'une théorie de changement

La présentation d'une ToC ne diffère pas de celle d'un modèle logique. Ajouter des hypothèses et des risques à l'outil visuel pourrait représenter un défi pour quiconque se lance dans l'élaboration d'un tel outil. Le risque d'alourdir la représentation visuelle peut nuire à tout l'intérêt de celle-ci, soit de « rendre l'information plus accessible et attirante » (Jones *et al.*, 2019). Pour pallier ces difficultés, il est fortement conseillé de produire trois versions de la théorie de changement :

- Un *texte explicatif* qui résume en quelques paragraphes l'ensemble de l'intervention et comment celle-ci est censée fonctionner. À ce texte explicatif, il est primordial de joindre l'énoncé des hypothèses et des risques. Cette version en forme de « synopsis » pourrait être utile au gestionnaire pour défendre l'intervention à des interlocuteurs externes. Ce texte narratif n'est pas nécessairement lié à la théorie du changement. Dans la pratique, il est toujours recommandé, voire exigé, d'avoir un écrit narratif pour expliquer comment l'ensemble du programme fonctionne. Ce texte accompagne toujours le modèle logique.
- Un *schéma* indiquant les liens causaux et les principales hypothèses qui montrent en un coup d'œil le fonctionnement de l'intervention. Cette version peut s'avérer utile pour tout public, incluant les partenaires externes et les bénéficiaires.
- Un *texte détaillé* qui explique en profondeur la théorie sous-jacente, les liens causaux et l'ensemble des hypothèses/risques de l'intervention. Ce texte détaillé peut servir aux fins de formation de nouveaux employés, en même temps qu'il sert de guide au gestionnaire et à son équipe de spécialistes (Mayne, 2017).

Pour faciliter l'élaboration de la version schématique de la ToC, Internet fournit une gamme intéressante de logiciels. Nous avons établi ici un tableau qui résume les principales caractéristiques de certains d'entre eux.

**Tableau 4** — Caractéristiques de logiciels pour l'élaboration d'une version schématique de la Toc. Inspiré de « Theory of Change Software » dans [BetterEvaluation.org](http://BetterEvaluation.org).

Logiciel	Caractéristiques principales
<b>Lucidchart</b>	<ul style="list-style-type: none"> <li>▪ Logiciel gratuit en ligne (ne requiert pas d'installation);</li> <li>▪ Interface facile d'utilisation (fonctionnalités de type « glisser-déposer »);</li> <li>▪ Intégration avec d'autres logiciels de productivité (Google Workspace, Confluence, Jira Software, YouTube, Facebook, Dropbox);</li> <li>▪ Téléchargement des schémas en format PDF, en format PNG, entre autres;</li> <li>▪ Partage facile.</li> </ul>
<b>Changeroo</b>	<ul style="list-style-type: none"> <li>▪ Logiciel en ligne (abonnement requis);</li> <li>▪ Orienté vers la cocréation des ToC complexes;</li> <li>▪ Interactif;</li> <li>▪ Utilisation simultanée avec plusieurs collaborateurs.</li> </ul>
<b>Theory Maker</b>	<ul style="list-style-type: none"> <li>▪ Logiciel gratuit en ligne;</li> <li>▪ Orienté vers la création des ToC et de cadres logiques;</li> <li>▪ L'utilisateur peut se concentrer sur le contenu : le logiciel crée le diagramme automatiquement.</li> </ul>
<b>Toco</b>	<ul style="list-style-type: none"> <li>▪ Logiciel en ligne (ne requiert pas de téléchargement);</li> <li>▪ Utile pour la création des ToC complexes;</li> <li>▪ Intégration des données avec Excel;</li> <li>▪ Partage facile avec plusieurs collaborateurs.</li> </ul>

Rappelons qu'il ne faudrait pas sous-estimer la présentation schématique de la ToC ou d'un modèle logique. Elle peut affecter la compréhension et la perception du programme par les spécialistes de l'évaluation et les parties prenantes. Selon Jones *et al.* (2019), les recherches sur les aspects cognitifs de la schématisation démontrent que des outils visuels contenant la même information peuvent être interprétés de façon différente. Un effort dans la présentation visuelle de ces outils facilite leur efficacité, permet d'obtenir un gain du temps de lecture et de compréhension et augmente leur crédibilité (Jones *et al.*, 2019, p. 14). Pour plus d'informations sur la façon d'améliorer la présentation d'une théorie de changement ou d'un modèle logique, le lecteur peut se référer aux recherches de Dilhon et Vaca (2018).

## Vérification de la qualité d'un modèle logique et d'une théorie de changement

Le tableau suivant énonce une série de questions de vérification du modèle logique et la théorie de changement. Des réponses affirmatives permettent de tester la qualité du modèle logique ou de la théorie de changement.

**Tableau 5** — Liste des aspects à prendre en compte pour un modèle logique/ToC de qualité.  
Inspiré de Porteous, 2013, p. 106.

Planification
<ul style="list-style-type: none"> <li><input type="checkbox"/> Tous les documents pertinents ont-ils été consultés avant la réalisation du modèle logique/ToC?</li> <li><input type="checkbox"/> Le modèle logique/ToC comporte-t-il une version simplifiée et une version exhaustive en annexe?</li> <li><input type="checkbox"/> Les bénéficiaires de l'intervention ont-ils été bien définis?</li> <li><input type="checkbox"/> Les différentes parties prenantes ont-elles été consultées?</li> </ul>
Élaboration et contenu
<ul style="list-style-type: none"> <li><input type="checkbox"/> Tous les éléments du modèle logique ont-ils clairement été énoncés?</li> <li><input type="checkbox"/> Les activités sont-elles formulées en verbes d'action?</li> <li><input type="checkbox"/> Les extrants ont-ils bien été formulés comme des produits découlant des activités?</li> <li><input type="checkbox"/> Les activités et résultats répondent-ils au besoin des groupes bénéficiaires du programme?</li> <li><input type="checkbox"/> Les différents résultats reflètent-ils les changements des niveaux différents?</li> <li><input type="checkbox"/> Les hypothèses et les risques ont-ils bien été formulés? Ont-ils été testés selon la logique « si...et...alors »?</li> <li><input type="checkbox"/> Les résultats peuvent-ils être mesurés à l'aide d'indicateurs?</li> <li><input type="checkbox"/> Les résultats reflètent-ils la direction, le changement escompté, les bénéficiaires et la portée de l'intervention?</li> </ul>
Processus participatif
<ul style="list-style-type: none"> <li><input type="checkbox"/> Le modèle logique/ToC a-t-il été réalisé selon un processus participatif?</li> <li><input type="checkbox"/> Les points de vue de toutes les parties ont-ils été pris en compte?</li> <li><input type="checkbox"/> Les parties prenantes sont-elles à l'aise avec la version finale du modèle logique/ToC?</li> </ul>
Présentation visuelle
<ul style="list-style-type: none"> <li><input type="checkbox"/> Le schéma tient-il sur une seule page?</li> <li><input type="checkbox"/> Le schéma peut-il être lu facilement et de façon rapide?</li> <li><input type="checkbox"/> La présentation visuelle est-elle attirante et crédible?</li> </ul>

## Conclusion

L'élaboration d'un modèle logique ou d'une théorie de changement est une étape fondamentale du processus d'évaluation d'une intervention. Si l'intervention à évaluer possède un modèle logique ou une ToC, les spécialistes de l'évaluation s'en servent pour comparer les impacts réels avec les résultats planifiés pour apporter un jugement de valeur au programme. À l'inverse, si le programme à évaluer ne possède pas un modèle logique ou une ToC, les spécialistes de l'évaluation entameront le processus de conception de cet outil avec les collaborateurs de l'organisation. Les avantages de cet exercice de modélisation sont multiples. Modéliser l'intervention permet la production d'outils de planification, de communication, de suivi et d'évaluation. Un bon modèle logique ou une bonne ToC est gage de solidité et de cohérence. Lorsque l'on explicite les enchaînements logiques et que l'on procède à une réflexion poussée des hypothèses et des risques, l'intervention aura plus de chances d'atteindre les changements de bien-être escomptés.

En outre, il ne faudrait pas que la personne chargée de l'évaluation s'engage dans ce processus de façon individuelle. La modélisation doit être réalisée idéalement dans le cadre d'un processus participatif. La consultation des intervenants, des partenaires et des bénéficiaires garantit une vision commune sur les changements à produire, et permet de canaliser les forces vers l'atteinte des objectifs. Le modèle logique et la théorie de changement facilitent le développement des indicateurs d'activités, des processus et des résultats. Ils constituent en définitive la base pour un examen des causalités des programmes en vue de leur amélioration continue (Mayne, 2017).

## Pour aller plus loin

Rogers, P. (2014). *Théorie du changement* (notes méthodologiques n° 2). UNICEF. <https://www.unicef-irc.org/publications/pdf/MB2FR.pdf>

Cette note méthodologique permet d'approfondir les connaissances sur l'utilisation des théories de changement en contexte international.

Funnell, S. C. et Rogers, P. J. (2011). *Purposeful program theory : Effective use of theories of change and logic models*. Jossey-Bass.

Ce livre explore davantage l'importance de la théorie de programme pour l'élaboration d'évaluations efficaces.

Secrétariat du Conseil du Trésor du Canada. (2019). *Intégration de l'analyse comparative entre les sexes plus dans l'évaluation : un guide d'introduction* (2019). <https://www.canada.ca/fr/secretariat-conseil-tresor/services/verifications-evaluations/evaluation-gouvernement-canada/gba-guide.html>

Le guide sur l'analyse comparative entre les sexes (ACS+) permet de comprendre la richesse et l'importance d'intégrer les facteurs identitaires dans l'ensemble des composantes de l'évaluation.

## Références bibliographiques

- Abdi, S. et Mensah, G. (2016). *Pleins feux sur : Modèle logique – Outil de planification et d'évaluation*. Santé publique Ontario. <https://www.publichealthontario.ca/-/media/documents/F/2016/focus-on-logic-model.pdf?la=fr>
- Alkin, M. C. et Taut, S. (2003). Unbundling Evaluation Use. *Studies in Educational Evaluation*, 29(1), 1-12. 10.1016/S0191-491X(03)90001-0
- Alkin, M. C. et King, J. A. (2016). The Historical Development of Evaluation Use. *American Journal of Evaluation*, 37(4), 568-579. <https://doi.org/10.1177/1098214016665164>
- Brousselle, A. et Buregeya, J. M. (2018). Theory-based evaluations : Framing the existence of a new theory in evaluation and the rise of the 5th generation. *Evaluation*, 24(2), 153-168. <https://doi.org/10.1177%2F1356389018765487>
- Chaumont, M. et Leroux, J. L. (2018). Le jugement évaluatif : subjectivité, biais cognitifs et postures du professeur. *Pédagogie collégiale*, 31(3), 27-33.
- Chen, H., Wendy Pan, H.-L., Morosanu, L. et Turner, N. (2018). Using Logic Models and the Action Model/Change Model Schema in Planning the Learning Community Program : A Comparative Case Study. *The Canadian Journal of Program Evaluation*, 33(1), 49-68. <https://doi.org/10.3138/cjpe.42116>
- Community Sustainability Engagement Evaluation Toolbox. (2010). *Program Logic*. [http://evaluationtoolbox.net.au/index.php?option=com\\_content&view=article&id=30&Itemid=136](http://evaluationtoolbox.net.au/index.php?option=com_content&view=article&id=30&Itemid=136)
- Dhillon, L. et Vaca, S. (2018). Refining theories of change. *Journal of Multidisciplinary Evaluation*, 14(30), 64-87.
- Douthwaite, B., Ahmad, F. et Shah, G.-H. (2020). Putting Theory of Change into Use in Complex Settings. *The Canadian Journal of Program Evaluation*, 35(1), 35-52. <https://doi.org/10.3138/cjpe.43168>
- Douthwaite, B., Mayne, J., McDougall, C. et Paz-Ybarnegaray, R. (2017). Evaluating complex interventions : A theory-driven realist-informed approach. *Evaluation*, 23(3), 294-311. <https://doi.org/10.1177%2F1356389017714382>
- Dumaine, F. (2012). Je me souviens... de t'avoir trop longtemps cherché. *The Canadian Journal of Program Evaluation*, 25(3), 95-102.
- Forss, K., Rebien, C. C. et Carlsson, J. (2002). Process Use of Evaluations : Types of Use that Precede Lessons Learned and Feedback. *Evaluation*, 8(1), 29-45.
- Funnell, S. C. et Rogers, P. J. (2011). *Purposeful program theory : effective use of theories of change and logic models* (1<sup>re</sup> édition). Jossey-Bass.
- Jabeen, S. (2018). Unintended outcomes evaluation approach : A plausible way to evaluate unintended outcomes of social development programmes. *Eval Program Plann.*, 68, 262-274. <https://doi.org/10.1016/j.evalprogplan.2017.09.005>
- Jones, N., Azzam, T. et Linnell Wanzer, D. (2019). Enhancing the Effectiveness of Logic Models. *American Journal of Evaluation*, 41(3), 452-470. <https://doi.org/10.1177%2F1098214018824417>



- Kinyuira, D. (2019). Using results chain framework as a tool for the improvement of performance evaluation in firms. *Journal of Strategy and Performance Management*, 7(1), 28-36.
- Lemon, A. et Pinet, M. (2018). Measuring unintended effects in peacebuilding : What the field of international cooperation can learn from innovative approaches shaped by complex contexts. *Eval Program Plann.*, 68, 253-261. <https://doi.org/10.1016/j.evalprogplan.2017.09.007>
- Lopez, J., Messomo Beyene, S. et El Kaddouri, F. (2020). *Intervention des renforcement des capacités évaluatives de la Direction des affaires internationales de l'Énap auprès d'un organisme au Maroc*. École nationale d'administration publique.
- Macfarlan, A. (2020). *Theory of Change Software*. Better Evaluation. [https://www.betterevaluation.org/en/evaluation-options/TOC\\_software](https://www.betterevaluation.org/en/evaluation-options/TOC_software)
- Mayne, J. (2017). Théories du changement : comment élaborer des modèles utiles. *La Revue canadienne d'évaluation de programme*, 32(2), 174-201. <https://doi.org/10.3138/cjpe.31144>
- Montague, S. (2019). Does your implementation fit your theory of change? *Canadian Journal of Program Evaluation*, 33(3), 316-335. <https://doi.org/10.3138/cjpe.53008>
- Patton, MQ. (2011). *Developmental evaluation : Applying complexity concepts to enhance innovation and use*. Guilford Press.
- Pawson, R. et Tilley, N. (1997). *Realistic evaluation*. SAGE Publications.
- Porteous, N. L. (2013). La construction du modèle logique d'un programme. Dans C. Dagenais et V. Ridde (dir.), *Approches et pratiques en évaluation de programmes*. Presses de l'Université de Montréal.
- Prinsen, G. et Nijhof, S. (2015). Between logframes and theory of change : reviewing debates and a practical experience. *Development in Practice*, 25(2), 234-246. <https://doi.org/10.1080/09614524.2015.1003532>
- Rey, L., Brousselle, A. et Dedobbeleer, N. (2012). Logic Analysis : Testing Program Theory to Better Evaluate Complex Interventions. *The Canadian Journal of Program Evaluation*, 26(3), 61-89.
- Ringhofer, L. et Kohlweg, K. (2019). Has the Theory of Change established itself as the better alternative to the Logical Framework Approach in development cooperation programmes? *Progress in Development Studies*, 19(2), 112-122. <https://doi.org/10.1177/1464993418822882>
- Rogers, P. J. (2016). *Théorie du changement*. UNICEF. <https://www.unicef-irc.org/publications/pdf/MB2FR.pdf>
- Secrétariat du Conseil du Trésor du Canada. (2012). *Approches d'évaluation axées sur la théorie : Concepts et pratiques*. <https://www.canada.ca/fr/secretariat-conseil-tresor/services/verifications-evaluations/evaluation-gouvernement-canada/approches-evaluation-axes-theorie-concepts-pratiques.html>
- Secrétariat du Conseil du Trésor du Canada. (2019). *Intégration de l'analyse comparative entre les sexes plus dans l'évaluation : un guide d'introduction (2019)*. <https://www.canada.ca/fr/secretariat-conseil-tresor/services/verifications-evaluations/evaluation-gouvernement-canada/gba-guide.html>



- Secrétariat du Conseil du trésor du Québec. (2018). *Cadre d'évaluation : Instructions destinées aux ministères et aux organismes*. [https://www.tresor.gouv.qc.ca/fileadmin/PDF/cadre\\_gestion/cadre\\_evaluation.pdf](https://www.tresor.gouv.qc.ca/fileadmin/PDF/cadre_gestion/cadre_evaluation.pdf)
- TIESS. (2018). *La Théorie du changement : fiche de synthèse*. [http://www.tiess.ca/wp-content/uploads/2018/04/TIESS\\_fiche\\_ToC\\_2018\\_04\\_10.pdf](http://www.tiess.ca/wp-content/uploads/2018/04/TIESS_fiche_ToC_2018_04_10.pdf)
- Tsala, U. B. (2019). La Chaîne des Résultats et la Théorie du Changement pour améliorer le cadre conceptuel d'une évaluation d'impact. *African Evaluation Journal*, 7(1), 1-12. <https://doi.org/10.4102/aej.v7i1.277>
- W.K. Kellogg Foundation. (2004). *Logic Model Development guide : Using Logic Models to Bring Together Planning, Evaluation and Action*. <https://ag.purdue.edu/extension/pdehs/Documents/Pub3669.pdf>

# Le processus d'évaluation

Richard Marceau



## Introduction

La modélisation de l'**intervention**<sup>1</sup>, telle que proposée au chapitre 8, vise à décrire l'intervention de manière structurée et détaillée. Concrètement, elle conduit à rassembler des informations et des observations souvent disparates qui, mises en forme, révèlent l'intervention dans sa nature propre et dans sa dynamique. C'est l'**objet d'évaluation** décrit dans sa structure fine.

Mais décrire ne suffit pas! Évaluer doit permettre en plus de dégager la **valeur de l'intervention**. L'évaluation doit interroger le **modèle de l'intervention**, plus exactement la représentation que l'on se fait de l'intervention, pour estimer sa capacité à produire une valeur sociale. Cette interrogation n'est pas un amalgame de préoccupations diverses et contextuelles : c'est une suite ordonnée d'études du modèle de l'intervention. Cette suite d'études, c'est le **processus d'évaluation** que l'on définit ainsi : « Processus visant à concrétiser une évaluation par la conduite d'une série d'études ».

Le présent chapitre propose un processus d'évaluation constitué d'études thématiques, desquelles découleront les grandes questions d'évaluation qui seront couvertes au chapitre 10. Ce processus d'évaluation est strictement d'ordre « analytique », c'est-à-dire qu'il est vu comme le déploiement d'un raisonnement évaluatif qui s'appuie autant sur des théories descriptives et des principes normatifs que sur des collectes de données et des observations de terrain. Les manières dont il peut être incarné par des institutions, par des **maîtres d'œuvre**, par des **évaluateurs/évaluatrices** et des **experts sectoriels/expertes sectorielles** ainsi que par des **parties prenantes**, sont dépendantes des différents contextes dans lesquels les interventions et les évaluations ont lieu. Cette préoccupation plus opérationnelle de l'évaluation est l'objet du chapitre 5 : *La gestion d'une évaluation*. Ce chapitre-ci pose les balises d'un processus nécessaire, donc d'une série d'étapes analytiques, pour que le travail de toutes les parties impliquées contribue à un exercice crédible et pertinent d'évaluation. Le rôle central des évaluateurs et évaluatrices dans une évaluation milite pour qu'ils comprennent en profondeur les exigences d'un processus systématique d'évaluation. Soulignons toutefois dès à présent que l'expertise sectorielle doit être sollicitée à des moments clés et fait partie intégrante du processus d'évaluation.

Plus spécifiquement, dans ce chapitre :

- Nous définissons la portée du processus d'évaluation;
- Nous soulignons l'importance qui doit être donnée au temps et au cycle de vie de l'intervention dans un processus d'évaluation;

---

1. Pour une définition du terme *intervention* ainsi que pour celles des autres termes d'évaluation marqués en gras à leur première occurrence dans ce chapitre, voir le glossaire du chapitre 2 : *La terminologie de l'évaluation*.

- Nous proposons des thématiques du modèle de l'intervention devant faire l'objet d'études particulières dans le processus d'évaluation;
- Nous composons trois cas types d'évaluation selon le moment de l'évaluation;
- Enfin, nous illustrons le processus d'évaluation à l'aide d'un cas d'initiative de développement.

Nous tissons ainsi des liens étroits et logiques entre le modèle de l'intervention (chapitre 8), le processus d'évaluation (ce chapitre) et les questions d'évaluation (chapitre 10). Il importe que les questions d'évaluation, qui orientent en dernière instance l'exercice concret d'évaluation, soient élaborées à partir d'un cadre rigoureux assurant un examen systématique et signifiant des interventions. Et ce cadre, c'est le processus d'évaluation.

## Évaluer l'intervention dans son entièreté

Les divers contextes d'intervention et d'évaluation invitent à poser des questions sur des éléments très précis d'une intervention. L'urgence des problèmes, le caractère dramatique des événements, le rapport de forces entre divers groupes d'intérêt ou les priorités politiques engendrent une série de préoccupations qui peuvent être de nature conjoncturelle. Celles et ceux impliqués dans une évaluation peuvent en conséquence faire l'objet de pressions les amenant à orienter le processus d'évaluation dans une direction ou dans une autre. Bien que l'évaluation se doit d'être sensible au contexte dans lequel elle se déroule, le processus d'évaluation a aussi ses propres exigences pour demeurer cohérent, objectif et signifiant pour la **société**.

Par exemple, pour des raisons liées aux cycles électoraux, le niveau politique peut être enclin à prioriser l'**étude de la mise en œuvre** de l'intervention plutôt que l'**étude des effets**, puisque les effets se produisent à plus long terme. Certains effets appréhendés d'une intervention peuvent mobiliser des groupes d'intérêt ou des communautés particulières pour que l'évaluation mette l'accent sur certaines études plutôt que sur d'autres. Les considérations d'équité peuvent amener à s'interroger prioritairement sur la représentation de divers groupes dans la **conception** de l'intervention et moins sur ses effets.

Les attentes face à l'évaluation sont multiples. Chacune de ces attentes ne recouvre à elle seule qu'une fraction des étapes à franchir pour aboutir à une évaluation rigoureuse et pertinente socialement. De plus, ces attentes s'expriment avec une puissance variable, selon les capacités sociales, économiques ou politiques des **acteurs**. Concentrer le processus d'évaluation sur les préoccupations des forces dominantes de la société devrait questionner : après tout, les interventions sujettes à évaluation ont souvent été conçues et appuyées par ces mêmes forces. Évaluer une intervention uniquement à partir des préoccupations de ceux et celles qui l'ont conçue et défendue créerait, paradoxalement, un profond biais, un conflit d'intérêts.

Aussi, il n'est pas concevable que le processus d'évaluation soit défini fondamentalement par un rapport de forces sociales, politiques ou économiques. La finalité même de l'évaluation est de prendre du recul sur ce qui a été fait pour éclairer la société, et non pas de reproduire les consensus initiaux à l'origine des interventions! S'il advenait que l'évaluation soit l'occasion de reproduire le statu quo, elle perdrait sa pertinence, et l'énergie consacrée serait dissipée inutilement. Le fait que l'évaluation soit consacrée à des clientèles, même diverses, ne doit pas rendre celles-ci maîtresses du processus d'évaluation. Par analogie, le patient, client de la pratique médicale, ne peut définir les procédures de diagnostic de ses propres problèmes de santé. Il peut, certes, choisir son médecin, mais pas le

processus médical ! Et puisque la profession médicale est réglementée, le processus médical n'est pas systématiquement l'objet de rapport de forces : il est basé sur la connaissance du corps humain et de son fonctionnement. Il en va de même pour l'évaluation, dont le processus doit être ancré dans la connaissance même de l'intervention et de son fonctionnement.

Pour éviter que le processus d'évaluation ne fasse l'objet de pressions indues qui l'éloignerait de sa finalité, il importe que sa définition même s'enracine dans l'objet d'évaluation, tout au moins dans la représentation que l'on s'en fait par le modèle de l'intervention. Le processus d'évaluation doit alors couvrir, dans son essence et dans le meilleur des cas bien sûr, l'ensemble des éléments du modèle de l'intervention et non seulement une de ses parties. Ainsi, le processus d'évaluation gagne : 1) en neutralité face aux acteurs les plus puissants; 2) en valeur sociale puisque les diverses facettes de l'évaluation ont leurs correspondances dans les attentes diversifiées de la société; 3) en cohérence et en rigueur par l'examen ordonné de l'intervention, de son origine à ses effets les plus fondamentaux.

Tout comme les étudiants en médecine qui se préparent à rencontrer une grande variété de problèmes de santé, les évaluateurs et évaluatrices doivent développer, comme porteurs du champ de connaissances et de pratique, leur expertise, leur maîtrise de la profession, en fonction d'une grande variété de situations. Et la somme de ces situations possibles, c'est l'entièreté du questionnement de l'intervention. Tout dans le modèle de l'intervention se prête et est susceptible de contribuer à l'évaluation, de la conception à la **performance** de l'intervention en passant par sa **mise en œuvre**. Par exemple, objecter que ce qui compte, ce sont surtout les **résultats**, ou encore les effets de l'intervention, c'est considérer qu'il n'y a pas d'après, qu'il n'y a que le temps présent qui se répète, indéfiniment. Qu'il n'y a pas de leçons à tirer d'un échec, que seul le succès nous apprend la manière d'agir, et que la formule gagnante doit être répliquée telle quelle pour produire systématiquement les mêmes effets.

C'est faire fi de l'évolution des contextes et des situations d'intervention. C'est faire fi des leçons à tirer d'un demi-échec ou même d'un échec complet. Le progrès se réalise en expliquant les erreurs autant que les bonnes performances. Et l'évaluation est là pour tirer des leçons, pour dégager une compréhension des succès et des échecs des interventions. Elle peut ainsi recommander avec pertinence des améliorations aux interventions en cours. Ainsi, l'évaluation devient un vecteur de changement, de progrès social. Pour cela, il faut apprendre des résultats des interventions, mais il faut aussi relier les résultats à leurs causes, savoir que les difficultés ou les succès des interventions proviennent des contextes d'intervention, de leur mise en œuvre ou de leur conception.

Il en découle que la meilleure stratégie d'apprentissage de l'évaluation est d'approcher l'intervention dans son entièreté. L'apprenant doit considérer l'intervention dans toute son ampleur et l'interroger aussi fondamentalement et largement que possible. L'évaluation prend donc naturellement la forme d'un large questionnement sur l'intervention, de sa naissance à son aboutissement, pour découvrir sa valeur sociale et les implications pour l'avenir.

Le processus d'évaluation devient ainsi un système d'examen de l'ensemble du modèle d'intervention et donc de ses parties. C'est de cette manière que l'évaluation peut se déployer et répondre le plus complètement possible aux besoins évaluatifs de la société et de ses acteurs. C'est en développant les compétences nécessaires à cette vision de l'évaluation que l'apprenant pourra afficher toute sa compétence professionnelle, quel que soit le contexte d'évaluation qui se présente à lui.

Les sections suivantes verront à prendre en compte certaines contraintes inhérentes à la pratique évaluative telles que le cycle de vie de l'intervention, le moment de l'évaluation et son contexte.

## Le processus d'évaluation, le temps et le cycle de vie de l'intervention

La logique de l'évaluation ne peut se déployer dans un univers abstrait et clos. Évaluer exige de comprendre l'objet d'évaluation, le contexte qui l'a vu naître, pourquoi et comment il est conçu et ce qu'il engendre. Le processus d'évaluation concrétise l'examen de l'intervention en s'imprégnant des caractéristiques et des propriétés de son objet d'évaluation.

Or, s'il est une chose sûre, c'est que l'intervention se déploie dans le temps. Forcément. L'intervention émerge d'un contexte particulier à un moment donné. Ce contexte est interprété par la société et par l'autorité qui conçoit et, par la suite, qui met en œuvre l'intervention. Plusieurs acteurs et actrices voient à sa mise en œuvre, ce qui nécessite temps et ressources pour produire des résultats qui n'apparaîtront également qu'au fil du temps.

Le processus d'évaluation ne peut faire fi de cette dynamique temporelle sans se révéler impertinent. Interpréter l'intervention ou un de ses éléments sans tenir compte de ce qui précède conduit à perdre le fil conducteur permettant la compréhension de l'intervention et de ses conséquences. La dynamique de l'intervention, son « cycle de vie », doit être totalement intégrée au processus d'évaluation pour interpréter l'intervention dans son contexte. Le processus d'évaluation doit intégrer l'« intelligence temporelle » de l'intervention.

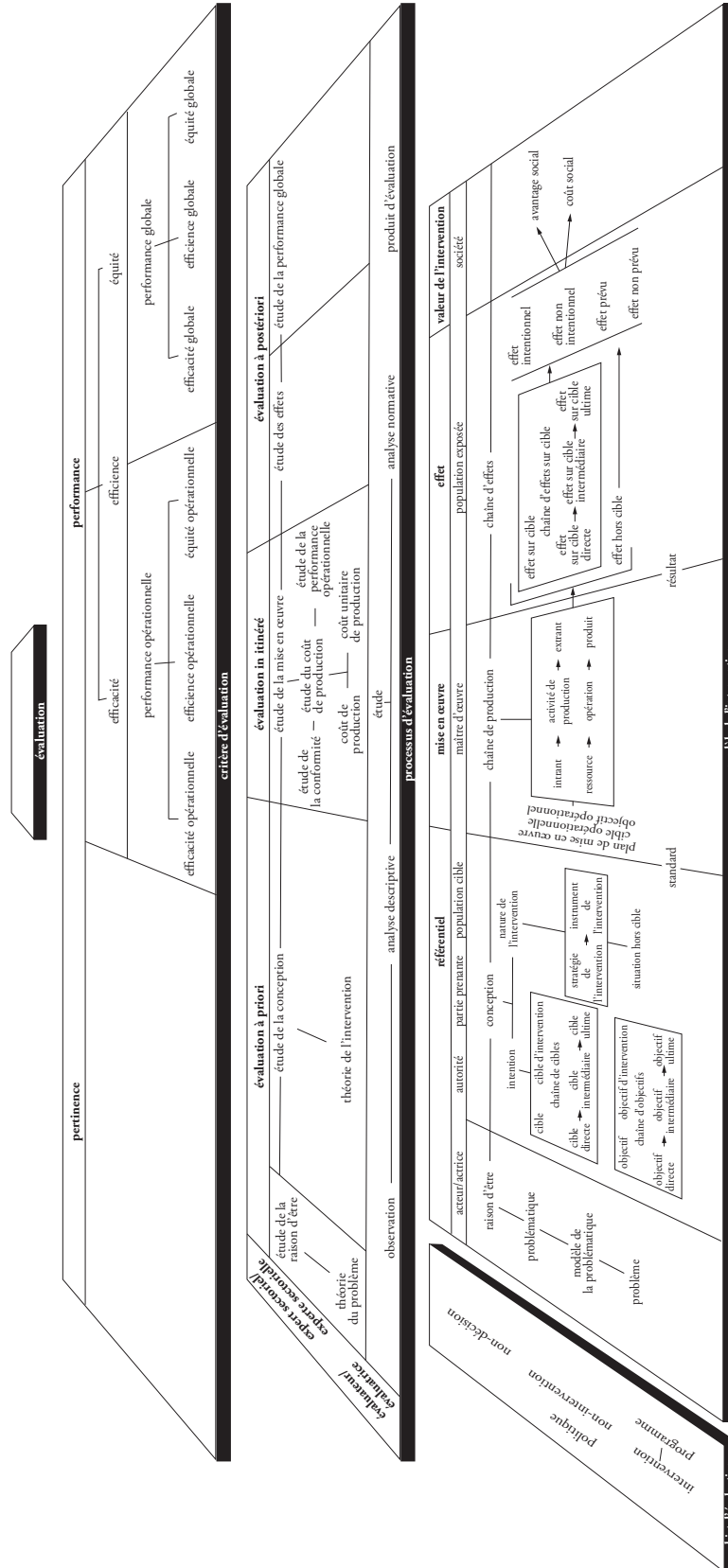
Le processus d'évaluation se doit alors d'accompagner l'intervention dans le temps. Il reconnaîtra ainsi le caractère fondamental du cycle de vie de l'intervention. L'examen de l'intervention tiendra compte des conditions associées au contexte initial d'intervention et à son évolution, jusqu'au moment de l'évaluation, de l'état des connaissances, des contraintes logistiques, sociales et politiques de la prise de décision, du contexte et du plan de mise en œuvre et des événements qui se déroulent parallèlement à la production des effets de l'intervention. Et tout cela dans un ordre donné, car la flèche du temps ne change jamais de direction ! Le processus d'évaluation doit accompagner l'intervention dans son cycle de vie ! L'intégration du processus d'évaluation à l'intervention elle-même doit être la plus complète possible tout en lui conservant l'indépendance politique nécessaire à son exécution.

Le tableau 1 illustre cet étroit assemblage temporel entre le modèle de l'intervention et le processus d'évaluation.

Le modèle d'intervention révèle l'écoulement du temps par l'explicitation du cycle de vie de l'intervention qui débute par la **problématique** de l'intervention et sa conception jusqu'à la production des effets, des **avantages sociaux** et des **coûts sociaux** de l'intervention (la représentation du flux du temps de gauche à droite n'est, bien entendu, qu'une convention et peut être réalisée de bien d'autres manières).

En conséquence, le processus d'évaluation, qui se superpose dans le tableau 1 au modèle d'intervention, doit synchroniser son examen sur le cycle de vie de l'intervention et le flux du temps. Nous verrons que les études proposées sont définies par ce que l'intervention offre à ce stade particulier de son évolution, permettant la prise en compte du contexte d'intervention et l'établissement des causes de sa performance ou de sa contre-performance.

Tableau 1 — Modèle d'intervention et processus d'évaluation



<https://enap.editionsjfd.com>

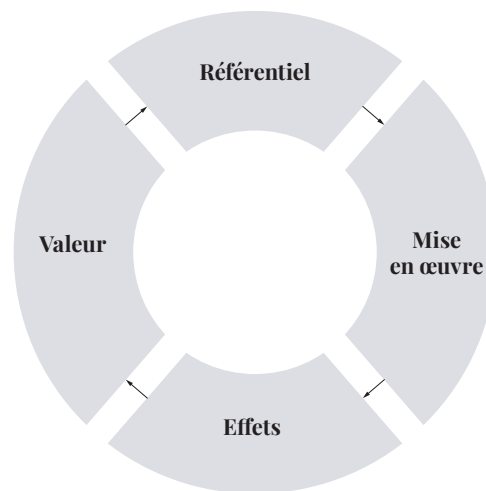
## Les études thématiques du processus d'évaluation

Dans l'ordre des choses, le processus d'évaluation morcelle et examine les quatre grandes dimensions de l'intervention telles que reflétées par le modèle d'intervention, pour mieux les intégrer par la suite dans l'évaluation<sup>2</sup> :

1. Le **référentiel**;
2. La **mise en œuvre**;
3. Les **effets**;
4. La **valeur de l'intervention**.

Ces dimensions reflètent naturellement le cycle de vie de l'intervention : la mise en œuvre est entreprise sur la base d'un référentiel; elle engendre les effets qui produisent la valeur de l'intervention.

Le fait que l'intervention puisse être appréhendée de manière dite « cyclique », c'est-à-dire qu'elle peut être revue, permet de profiter d'une mise en œuvre corrigée. Et le fait qu'elle enclenche ainsi de nouveaux effets, cette vision en boucle ne nie pas la vision temporelle, séquentielle et forcément causale du cycle de vie d'une intervention; c'est sa mécanique naturelle qui, au fil du temps, produit une histoire d'apparence cyclique, une histoire plus complexe de l'intervention. Cette manière de lire l'intervention est d'ailleurs en harmonie avec celle de l'« analyse de politiques publiques », champ de connaissances issu de la science politique et connexe à l'évaluation des politiques publiques<sup>3</sup>. On retrouve des formes analogues de représentation en évaluation du développement, par exemple chez Goldman et Pabari (2021), qui identifient dans leur cycle de politique et programme les étapes (1) de diagnostic; (2) de conception (référentiel); (3) de mise en œuvre; (4) des « impacts » et « répercussions » (effets).



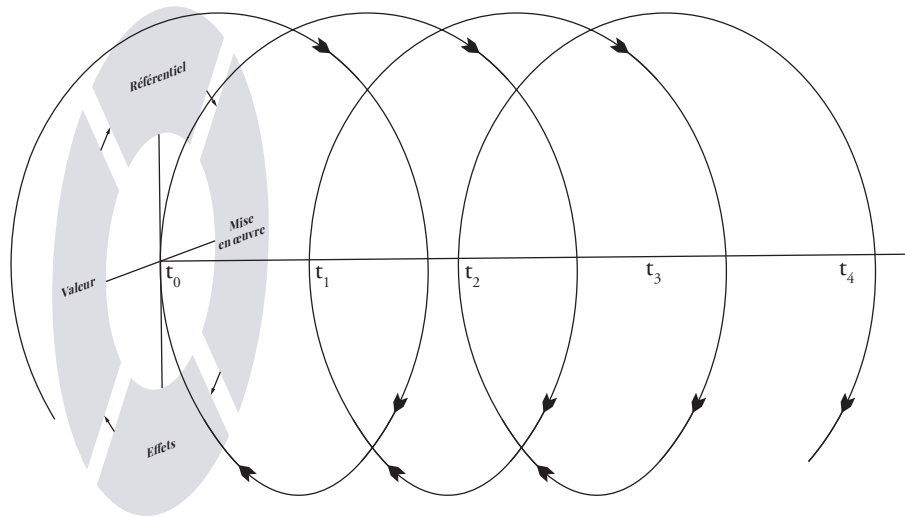
**Figure 1**

Représentation cyclique de l'évolution d'une intervention

2. Voir le *Dictionnaire terminologique de l'évaluation*, 2014.

3. Voir la bibliographie commentée.

Pour une compréhension plus entière de l'intervention et de son évaluation, il y a lieu de se représenter leur évolution d'une façon dynamique, comme l'illustre la figure 2. On y voit que le référentiel initial peut être revu dans le temps (au temps 1 ( $t_1$ ), temps 2 ( $t_2$ )...) grâce à l'apport de la connaissance de la mise en œuvre, de celle des effets et de la valeur de l'intervention mise en œuvre au temps 0 ( $t_0$ ). Ce cycle de vie d'apprentissage n'est pas à l'image d'une roue en perpétuel retour sur elle-même, mais reflète plutôt une trajectoire nourrie par le passage du temps, par son passé et son présent. L'évaluation peut contribuer à l'amélioration des interventions en continu ou à des moments précis.



**Figure 2**

Représentation dynamique de l'évolution d'une intervention

Le processus d'évaluation concrétisera l'évaluation en examinant chacune de ces quatre dimensions de l'intervention de manière séquentielle, en suivant la flèche du temps, et en réalisant des études dédiées à plusieurs thématiques. Voici notre proposition de huit études regroupées en cinq différentes thématiques du processus d'évaluation<sup>4</sup> :

**Tableau 2** — Les dimensions du modèle d'intervention et les études thématiques.

Dimensions du modèle d'intervention	Thématiques et études
Référentiel	1. Étude de la raison d'être 2. Étude de la conception
Mise en œuvre	3. Étude de la mise en œuvre Étude de la conformité Étude des coûts Étude de la performance opérationnelle
Effets	4. Étude des effets
Valeur de l'intervention	5. Étude de la valeur de l'intervention

4. Voir le *Dictionnaire terminologique de l'évaluation* (2014).



Cinq grandes études thématiques composent le processus d'évaluation. Chacune de ces études consiste à examiner une dimension de l'intervention correspondant à un ou des éléments particuliers du modèle d'intervention. Le **référentiel** de l'intervention est couvert par deux grandes études thématiques : l'**étude de la raison d'être** et l'**étude de la conception**. L'**étude de la mise en œuvre** est appuyée de trois études plus spécifiques, complémentaires et incontournables selon les exigences habituelles de la pratique administrative, soit l'**étude de la conformité**, l'**étude des coûts** et l'**étude de la performance opérationnelle**. L'**étude des effets** exige à elle seule une partie significative des ressources consacrées à l'évaluation. Enfin, l'**étude de la valeur de l'intervention** résume et conclut l'évaluation par des analyses descriptives et normatives de l'intervention.

Chaque étude contribue de manière unique à l'évaluation, en délimitant et en posant les questions qui découlent de sa thématique particulière et en apportant des réponses. C'est donc la thématique de l'étude qui détermine la nature et le contenu des questions posées. Les études thématiques concentrent les questions sur des éléments particuliers de l'intervention, ce qui évite de multiplier indûment les questions d'évaluation ou encore d'en oublier d'importantes. En couvrant ainsi le modèle d'intervention par cinq grandes études thématiques, il devient plus évident d'équilibrer la répartition, le nombre et le contenu des questions d'évaluation et, conséquemment, la répartition de l'effort consacré à l'évaluation.

Par exemple, l'étude de la raison d'être implique de poser les questions dont les réponses permettront de faire le tour de la thématique « raison d'être » de l'intervention. Il s'agira d'abord de poser des questions sur la raison d'être telle qu'interprétée par l'autorité responsable de l'intervention. L'étude impliquera aussi des questions et des sous-questions sur la manière dont la raison d'être a été élaborée, a été appuyée et fondée sur des observations, sur les diagnostics posés, sur les motifs qui légitiment l'intervention, etc. Au sortir de l'étude de la raison d'être, le processus d'évaluation a cerné la raison d'être telle qu'exprimée par l'autorité, mais a aussi développé une vision critique qui nourrira l'évaluation de l'intervention. L'évaluation répond ainsi à la question principale de l'étude thématique et à des sous-questions jugées essentielles à la finalité de l'étude. Autre exemple, l'étude de la conformité implique des questions sur la nature et sur le contenu du **plan de mise en œuvre**, mais aussi sur l'adéquation entre la mise en œuvre de l'intervention et le plan de mise en œuvre.

Chaque étude thématique a comme finalité de poser les bonnes questions et de trouver les réponses de manière à ce que, lorsque les différentes études sont complétées, les grandes dimensions de l'intervention soient couvertes par l'évaluation. Les études thématiques sont donc plus qu'un agrégat de questions : la thématique qui leur est confiée détermine les questions posées par les études. L'ensemble des questions posées par les études thématiques correspondent aux questions posées par le processus d'évaluation.

Ainsi se tisse le lien logique entre les questions d'évaluation, les études thématiques, le processus d'évaluation et l'évaluation elle-même. Le lien se construit grâce aux relations obligées qu'entretiennent les questions d'évaluation avec les études thématiques et grâce aux liens de dépendance qu'entretiennent les études thématiques avec le processus d'évaluation. Le pont entre modèle de l'intervention et processus d'évaluation se construit par la superposition naturelle du processus d'évaluation au modèle de l'intervention.

Il importe de préciser qu'une étude thématique n'est pas en soi une évaluation. Une étude des effets n'est pas une évaluation de l'intervention, car son spectre est trop restreint pour prétendre poser un jugement équilibré sur la valeur de l'intervention. Ainsi, le fait que les effets correspondent à des **effets intentionnels** ne permet pas de conclure que l'intervention est efficace. Des effets décevants ne signifient pas non plus une mise en œuvre défailante ou une mauvaise conception. Il faut en savoir plus. Enfin, une étude des effets n'est pas une évaluation des effets : attribuer une valeur sociale à des observations sur les effets nécessite un retour au référentiel.

Ceci montre que les questions d'évaluation et les études thématiques ne sont pas indépendantes les unes des autres. Les questions doivent être posées et les études doivent être réalisées dans un ordre donné, correspondant naturellement au cycle de vie de l'intervention. Les réponses apportées et les interprétations que l'on peut tirer d'une étude thématique doivent faire appel à d'autres études et à d'autres questions et réponses du processus d'évaluation. Les questions s'imbriquent les unes dans les autres.

Les études thématiques et leurs questions ont également un « moment propice ». L'étude de la mise en œuvre ne peut logiquement ni temporellement précéder l'étude de la raison d'être. L'ensemble de ces études et questions, l'ordre dans lequel se déroulent ces études et dans lequel les questions sont posées, les relations entre ces études et celles entre les questions, cet ensemble, c'est, bien sûr, le processus d'évaluation.

## Le moment de l'évaluation

Malgré la logique implacable du temps que nécessite la mise en œuvre et les effets des interventions, les nécessités de l'administration peuvent imposer leurs propres moments pour évaluer. Il arrive que l'on souhaite faire le point sur une intervention, même l'évaluer, avant que la mise en œuvre ne soit complétée; ou avant que les effets ne se soient tous produits. Ces besoins administratifs peuvent être légitimes, l'administration ayant sa propre logique. Elle force dans ces cas les responsables de l'évaluation à se compromettre prématurément sur la valeur d'une intervention. La responsabilité de l'équipe d'évaluation est alors de faire ce qui est possible de faire au moment où l'évaluation est demandée et d'exprimer les forces, mais aussi les faiblesses d'une telle évaluation réalisée à ce moment précis.

Les contraintes sur le moment de l'évaluation imposées par l'administration conduisent à définir des cas types d'évaluation. Nous dégagons ainsi trois cas types d'évaluation selon leur moment de réalisation :

1. L'évaluation **à priori**, celle qui est réalisée avant la mise en œuvre;
2. L'évaluation **in itinéré**, celle qui est réalisée pendant la mise en œuvre;
3. L'évaluation **à postériori**, celle qui est réalisée après la mise en œuvre.

Chaque type d'évaluation, exécutée à un moment particulier, autorise certaines études et questions, et en interdit d'autres.

**Tableau 3** — Les études thématiques et les cas types d'évaluation selon le moment d'évaluation.

Études thématiques	Cas types d'évaluation selon le moment d'évaluation
Étude de la raison d'être	Évaluation à priori
Étude de la conception	
Étude de la mise en œuvre	Évaluation in itinéré
Étude de la conformité	
Étude des coûts	
Étude de la performance opérationnelle	
Étude des effets	Évaluation à postériori
Étude de la valeur de l'intervention	

Ainsi, l'évaluation à priori permet l'étude de la raison d'être et l'étude de la conception (voir tableau 3). Elle exclut logiquement l'étude de la mise en œuvre et l'étude des effets. À l'administration qui désire connaître les effets d'une intervention avant sa mise en œuvre, il lui faut puiser dans l'étude de la conception pour connaître les **effets prévus** de l'intervention. Mais l'administration ne pourra évidemment obtenir les évidences factuelles que permet l'étude des effets, celle-ci ne pouvant forcément être exécutée qu'après l'apparition de ces effets.

En plus de bénéficier de l'étude de la raison d'être et de l'étude de la conception, l'évaluation in itinéré peut compter sur l'étude de la mise en œuvre et sur ses études spécifiques que sont l'étude de la conformité, l'étude des coûts et l'étude de la performance opérationnelle. Cependant, l'étude des effets et l'étude de la valeur de l'intervention lui sont logiquement inaccessibles.

C'est l'évaluation à postériori qui, seule, peut profiter de l'ensemble des cinq études thématiques, car son moment lui permet naturellement d'y accéder.

Puisque le processus d'évaluation vise l'entière du modèle d'intervention, il est correct, dans une logique temporelle ou séquentielle, d'affirmer que la somme des évaluations à priori, in itinéré et à postériori correspond à l'ensemble du processus d'évaluation. Avec du recul, cette trilogie séquentielle de l'évaluation reflète, tel un miroir, la trilogie « réflexion-action-réaction » du modèle de l'intervention, puisque l'évaluation à priori se questionne sur la réflexion à propos de l'intervention, l'évaluation in itinéré se concentre sur l'intervention elle-même et l'évaluation à postériori s'interroge sur la réaction à l'intervention.

## Une illustration du processus d'évaluation

Un cas fictif mais pouvant être considéré typique de plusieurs interventions en matière de lutte contre la pauvreté sera utilisé dans ce chapitre pour illustrer le déploiement d'un processus d'évaluation et les enjeux qu'il pose. Seulement certains éléments d'information qui sont nécessaires à ce chapitre seront précisés ici<sup>5</sup>.

« Une intervention en matière de lutte contre la pauvreté : cas fictif » (Vénétia Sauvain, École nationale d'administration publique, 2020, adapté de l'Initiative nationale pour le développement humain du Royaume du Maroc).

### 1. Contexte du cas :

*La prise de conscience d'une problématique de pauvreté, s'exprimant souvent sous la forme de précarité, d'exclusion sociale, de chômage, de délinquance et de misère, et touchant particulièrement les femmes et les jeunes, amène l'État à intervenir pour améliorer significativement la situation. Les niveaux ministériel, régional et local sont amenés à agir avec l'intention de diminuer l'ampleur de la pauvreté dans le pays. L'intervention vise à financer des projets d'activité génératrice de revenus (AGR) et de microprojets coopératifs (MPC). Un système de suivi-évaluation de l'intervention est mis en place avec la collaboration des différents paliers d'intervention.*

### 2. La portée du processus d'évaluation

La situation de pauvreté présente inévitablement un caractère dramatique à l'échelle humaine, et prend des proportions d'autant plus difficiles à soutenir lorsque l'on prend en compte l'ampleur de la population qui la vit au quotidien. De nombreux acteurs sociaux et institutionnels sont à la fois impliqués dans la réalité de la pauvreté et dans la recherche et la mise en œuvre de solutions. Chacune de ces parties prenantes possède sa propre compréhension de la situation, son interprétation des causes, son choix de la meilleure solution, son analyse de la manière de mettre en œuvre une solution, son anticipation des effets des solutions et même sa propre lecture des conséquences observées d'une intervention. Ces parties prenantes ont souvent des visions très variées des priorités à donner à un exercice d'évaluation. Certaines privilégieront l'étude de la mise en œuvre des différents paliers, d'autres préféreront que l'on observe sur le terrain les résultats concrets des interventions sur les conditions de vie des populations ciblées, d'autres s'interrogeront prioritairement sur le mécanisme de décision et les parties impliquées lors du choix des solutions au problème de pauvreté.

Chacune de ces parties prenantes exprime à sa manière une partie des préoccupations légitimes face à une intervention de lutte à la pauvreté. Mais ce qui doit retenir l'attention des évaluatrices et évaluateurs, c'est que toutes ces préoccupations sont liées les unes aux autres. Le succès de la mise en œuvre dépend en partie de la nature de l'intervention retenue; les résultats sur le terrain dépendent autant de la solution retenue que de sa mise en œuvre. Et l'efficacité de toutes les étapes de l'intervention dépend au préalable de la qualité du diagnostic du problème de pauvreté tel que vécu par les populations affectées.

5. Les chapitres portant sur les questions d'évaluation et la matrice d'évaluation détaillent plus encore comment une évaluation de ce cas peut être réalisée.

Prioriser au départ une thématique du processus d'évaluation, c'est le morceler et, en conséquence, réduire sa capacité à diagnostiquer les causes et les conséquences des effets d'une intervention, d'une initiative de développement. À moins que le processus d'évaluation ait été mis en place depuis plus d'un cycle de vie de l'initiative, et que la familiarisation de l'équipe d'évaluation puisse compter sur une riche expérience de la situation, il importera de veiller à ce que la portée de l'évaluation vise l'entièreté de l'initiative, c'est-à-dire à partir des diagnostics initiaux sur la problématique vécue sur le terrain, à la détermination des intentions, à l'examen des options d'intervention, au plan de mise en œuvre, aux effets à court, moyen et long terme, jusqu'à l'estimation des avantages et des coûts sociaux de l'initiative qui révéleront empiriquement sa valeur.

Ainsi seront respectées la cohérence et l'efficacité du processus évaluatif. Ainsi sera-t-il possible de répondre aux préoccupations des nombreuses parties prenantes de l'initiative et de son évaluation, et non seulement aux parties les plus puissantes. L'objectivité nécessaire à l'examen de la situation de pauvreté et à celui de l'efficacité et de l'efficience de l'intervention sera assurée et permettra de juger, de la manière la plus impartiale, la qualité de l'initiative. L'impartialité du processus d'évaluation n'est pas qu'un principe : c'est une condition de justesse du diagnostic et de l'amélioration rapide de la situation problématique de pauvreté.

### 3. L'importance du temps et du cycle de vie de l'intervention

Si ce drame a au moins comme contrepartie la stimulation de la volonté d'agir de l'État, le sentiment d'urgence qui s'en dégage pourrait inviter à prendre des décisions sur le processus d'évaluation aux effets contre-intuitifs.

Par exemple, si l'intervention n'atteint pas rapidement ses objectifs de réduction de la pauvreté, il pourrait apparaître logique d'arrêter l'intervention et d'y substituer une autre formule d'intervention préalablement reléguée au second plan lors de la conception de l'initiative. Le danger est possiblement de rejeter trop rapidement une option de solution qui n'a pas eu l'occasion de révéler toute son efficience.

L'initiative de développement doit avoir l'occasion de se déployer dans le temps pour manifester sa potentialité. L'évaluation doit s'ancrer dans la nature de l'intervention pour en apprendre sur sa performance véritable. Elle doit rythmer son processus selon le cycle de vie de l'initiative.

### 4. Les thématiques du modèle de l'intervention et les études en découlant

À ce stade du cas, il appert que les responsables de l'initiative sont en mesure de concevoir et de déployer un système de suivi-évaluation en collaboration avec les instances impliquées. Pour en arriver là, il importe que la raison d'être de l'intervention ait été réfléchi et énoncée, et que la conception de l'intervention soit une étape aboutie. L'énoncé du cas laisse supposer que la réflexion sur la problématique d'intervention a eu lieu (énoncée par ses multiples formes de *précarité, d'exclusion sociale, de chômage, de délinquance et de misère, et touchant particulièrement les femmes et les jeunes*). Elle laisse supposer également que l'**intention** de l'**autorité** ait été dûment exprimée (la *diminution de la pauvreté dans le pays*). Suivra la mise en œuvre de l'intervention et l'apparition des premiers effets. Le processus d'évaluation peut donc être immédiatement enclenché grâce à l'étude de la raison d'être et également grâce à l'étude de la conception de l'intervention. Nous verrons dans les prochains chapitres les questions d'évaluation qui pourront alors être posées et les exigences de collecte d'information qui en découlent pour les opérations de suivi nourrissant l'évaluation.

L'étude de la mise en œuvre pourra débiter lorsque le plan de mise en œuvre impliquant les diverses instances aura été élaboré, et que les premiers éléments de la « **chaîne de production** » apparaîtront. La production des premiers effets invitera ensuite à l'étude des effets.

### 5. Les conséquences du moment de l'évaluation

L'ambition de l'évaluation d'une intervention sur une problématique de pauvreté doit être en synchronie avec le moment auquel elle est enclenchée. Dans le cas de figure, l'ambition évaluative doit se limiter, à ce stade, à une évaluation à priori, qui autorise seulement l'étude de la raison d'être et l'étude de la conception. Ces deux études permettront de décrire et de juger de la valeur de l'intervention strictement quant à son référentiel.

Cette restriction peut sembler exagérée si l'on pense par exemple aux avis d'experts sollicités fréquemment sur la valeur de l'intervention avant même sa mise en œuvre. Ces avis ont, bien entendu, une grande valeur, puisqu'ils reposent généralement sur des connaissances et des expériences pertinentes quant aux problématiques impliquées et aux instruments d'intervention retenus. Mais il faut considérer qu'il s'agit forcément d'avis « à priori » reposant sur la performance anticipée des **instruments d'intervention**. Si ces avis ont nourri la réflexion sur la conception de l'intervention, ils peuvent renforcer la qualité de la conception de l'intervention.

Mais les limites de l'évaluation à priori doivent être comprises. Les meilleurs avis à priori ne peuvent se substituer totalement à une évaluation in itinéré, et encore moins à une évaluation à postériori. Seules les contraintes temporelles de l'administration ou du niveau politique justifient la réalisation d'une évaluation à priori. Néanmoins, l'étude de la raison d'être et l'étude de la conception, idéalement entreprises tôt dans le cycle de vie de l'initiative, constitueront un apport appréciable pour l'évaluation in itinéré et l'évaluation à priori.

## Bibliographie commentée

Le processus d'évaluation proposé ici est inspiré de l'approche de la *Public Policy Analysis*, initiée entre autres par Harold Lasswell (1956), Charles O. Jones (1970) et James E. Anderson (1975). Cette approche est centrée sur l'idée d'un cycle de politiques publiques (*Policy Cycle*) comprenant différents stades d'évolution. Dans une formule qui deviendra classique, Anderson a proposé un cycle de politiques publiques en cinq stades : 1) mise à l'agenda/identification du problème; 2) formulation de politiques; (3) décision; (4) mise en œuvre; 5) évaluation. Il s'agit d'une orientation « problème-solution » qui ne fait pas l'unanimité en sciences politiques à cause de ses préoccupations normatives, mais qui, néanmoins, mettra souvent l'accent sur les forces politiques en jeu à chaque stade du cycle de politiques publiques, soit l'analyse descriptive des politiques publiques. Par la suite, des auteurs tels Thomas Dye (1978) et William Dunn (2017) ont fait de manière plus marquée la transition vers l'analyse normative des politiques.

Le processus d'évaluation proposé ici découle de cette vision normative de l'analyse de politiques publiques. Le déploiement dans le temps de la vie d'une politique publique conduit à identifier un problème qui justifie une intervention de l'État. L'analyse se pose ensuite sur les options disponibles pour régler le problème, sur le processus permettant l'adoption de la solution, sur sa mise en œuvre et enfin sur le processus d'évaluation. Cet ancrage de l'évaluation dans l'analyse normative des politiques publiques permet d'élaborer un cadre global et séquentiel d'évaluation des interventions publiques comme celui présenté dans ce chapitre. Cette vue en séquences de l'intervention de l'État

est conforme avec l'idée du « *Logic Model* », centré sur la mise en œuvre et les effets, et qui se présente également sous forme de séquences : intrants – activité de production – extrant – effets. L'approche du cycle de politiques publiques englobe naturellement le « *Logic Model* » et permet l'élaboration du cadre d'évaluation avec plus de recul et plus cohérent, en balayant l'attention du problème jusqu'à son éventuelle résolution et en situant le tout dans son contexte politico-administratif. Des auteurs récents en évaluation (par exemple Goldman et Pabari (2017)) recourent d'ailleurs à cette vision de cycle de politique pour déployer leur approche d'évaluation.

## Messages clés

1. Le processus d'évaluation doit non seulement permettre la description de l'intervention mais aussi conduire à en estimer la valeur.
2. De nombreux acteurs sont impliqués dans le processus d'évaluation. Les évaluateurs et les évaluatrices, par leur rôle d'acteur et de facilitateur dans ce processus, doivent maîtriser la série d'étapes analytiques qu'exige une évaluation d'intervention.
3. La portée du processus d'évaluation couvre potentiellement tout le cycle de vie de l'intervention, de la raison d'être à la valeur de l'intervention.
4. Pour chaque grande dimension du modèle de l'intervention – référentiel, mise en œuvre, effets, valeur de l'intervention – correspond une ou des études thématiques du processus d'évaluation.
5. Le processus d'évaluation définit un cadre à l'intérieur duquel les questions d'évaluation peuvent être conçues, priorisées et ordonnées.
6. Trois cas types d'évaluation se présentent selon le moment de l'évaluation : l'évaluation à priori avant la mise en œuvre, l'évaluation in itinéré pendant la mise en œuvre, l'évaluation à posteriori après la mise en œuvre. Les études que chaque cas type d'évaluation permet sont déterminées par le cycle de vie de l'intervention.

## Exercice de réflexion

Le mandat d'évaluation qui est confié à votre équipe d'évaluation consiste à estimer les effets d'une initiative majeure de développement visant l'approvisionnement en eau potable d'une partie significative de la population du pays ainsi qu'à rendre compte de la valeur de l'intervention.

1. À quel cas type d'évaluation a-t-on affaire ?
2. Comment votre équipe procédera-t-elle pour planifier la mesure des effets ? Comment allez-vous identifier les effets potentiels de l'initiative pour bien organiser les mesures sur le terrain ? À quelle source de renseignements allez-vous puiser si vous ne détenez pas d'expertise particulière dans le domaine de l'approvisionnement en eau potable ?
3. Listez, à votre connaissance, les catégories d'acteurs avec lesquels votre équipe d'évaluation devra interagir ou sera confrontée au cours de l'évaluation. Associez à ces catégories d'acteurs des positions plausibles qu'ils adopteront lors de vos rencontres quant aux types d'effets à tenir compte dans votre évaluation. Vous attendez-vous à des points de vue convergents ou divergents ? Pourquoi y aurait-il convergence ? Ou divergence ?



4. Distinguez les types d'acteurs qui auront de la facilité à exprimer leurs points de vue et ceux qui auront plus de difficultés. Si les points de vue de ces acteurs divergent, que ferez-vous? Entre des rapports d'expertise bien rédigés et des points de vue communautaires spontanés ou moins structurés en termes d'argumentaire, comment arbitrez-vous ces points de vue sur les priorités de mesures des effets?
5. Comment décririez-vous votre rôle face à ces acteurs qui montrent un intérêt vif pour le mandat d'évaluation mais qui manifestent des positions différentes?
6. Même si le mandat d'évaluation porte exclusivement sur les effets et la valeur de l'intervention, se peut-il que des divergences s'observent chez les parties prenantes bien organisées ou moins bien organisées, ou encore du côté de l'autorité ayant confié le mandat quant à la nature ou l'ampleur du problème d'approvisionnement en eau potable? Si la divergence est sérieuse, considérez-vous que vous devrez aborder dans vos travaux et votre rapport d'évaluation ces divergences et votre analyse de la problématique d'intervention?
7. Si l'équipe d'évaluation est insatisfaite des renseignements qu'elle détient sur le référentiel de l'initiative et sur sa mise en œuvre, comment procéderait-elle pour justifier une demande de renseignements supplémentaires au mandant ou, plus encore, une extension du mandat permettant de couvrir les quatre grandes dimensions de l'initiative : le référentiel, la mise en œuvre, les effets et la valeur de l'intervention?
8. Dans le cas de figure où le mandant refuserait d'ouvrir le mandat sur le référentiel et la mise en œuvre, que feriez-vous pour remplir au mieux votre mandat?

## Pour aller plus loin

Pour en savoir plus sur l'analyse de politiques publiques et avoir accès à une documentation élaborée, voir l'entrée « politique publique » du *Dictionnaire encyclopédique de l'administration publique* de l'ENAP par Jean Turgeon et Jean-François Savard, ainsi que l'entrée « analyse de politiques (programmes) et cadre d'analyse » par Éric Montpetit.

[https://dictionnaire.enap.ca/Dictionnaire/19/Index\\_par\\_thematique.enap?by=them&id=14](https://dictionnaire.enap.ca/Dictionnaire/19/Index_par_thematique.enap?by=them&id=14)

[https://dictionnaire.enap.ca/Dictionnaire/19/Index\\_par\\_thematique.enap?by=them&id=12](https://dictionnaire.enap.ca/Dictionnaire/19/Index_par_thematique.enap?by=them&id=12)

Pour connaître le champ de l'évaluation de programme (*Program Evaluation*) et avoir accès à une documentation élaborée, voir l'entrée « évaluation de programme » du *Dictionnaire encyclopédique de l'administration publique* de l'ENAP rédigée par l'auteur de ce chapitre.

[https://dictionnaire.enap.ca/Dictionnaire/19/Index\\_par\\_thematique.enap?by=them&id=44](https://dictionnaire.enap.ca/Dictionnaire/19/Index_par_thematique.enap?by=them&id=44)



## Références bibliographiques

Anderson, J. E. (1975). *Public Policy Making*. Holt, Rinehart & Winston.

Dunn, W. N. (2017). *Public Policy Analysis. An Integrated Approach* (6<sup>e</sup> éd.). Routledge.

Dye, T. R. (1978). *Understanding Public Policy*. Engelwood Cliffs, Prentice-Hall.

Goldman, I. et Pabari, M. (2021). *Using Evidence in Policy and Practice : Lessons from Africa*. Routledge.

Jones, C. O. (1970). *An Introduction of the Study of Public Policy*. Belmont, Wadsworth Publishing Company.

Lasswell, H. D. (1956). *The Decision Process*. College Park, University of Maryland Press.

Marceau, R. et Sylvain, F. (2014). *Dictionnaire terminologique de l'évaluation : politiques, programmes, interventions : la dimension conceptuelle*. Les Éditions GID.

# Les questions d'évaluation

*Richard Marceau*



## Introduction

Il est fréquent qu'une équipe d'évaluation soit interpellée par toutes sortes de demandes et de préoccupations. Une intervention implique plusieurs parties prenantes et chacune se sent concernée à sa manière, en fonction de ses intérêts particuliers. Ces préoccupations pour la conduite d'une évaluation prennent généralement la forme de questions dont les réponses devraient en principe éclairer la prise de décision. La quantité de questions posées peut être inquiétante compte tenu des ressources toujours limitées pour l'évaluation; leur nature souvent disparate pose un enjeu de cohérence et d'exhaustivité pour l'évaluation, auquel il importe de répondre adéquatement. Enfin, les évaluateurs et évaluatrices, de par leurs compétences et leur expérience, sont naturellement amenés à analyser l'intervention sous des angles peut-être moins contextuels, mais tout aussi pertinents.

La conduite d'une évaluation passe alors par la priorisation et l'organisation des questions pour conférer à l'évaluation des qualités de cohérence, d'exhaustivité, de réalisme et d'efficacité.

Le chapitre 9 portant sur le processus d'évaluation a déjà proposé un cadre à l'intérieur duquel les questions d'évaluation peuvent être conçues et reçues, priorisées et ordonnées. Le processus d'évaluation définit la portée de l'évaluation, respecte la dimension temporelle et le cycle de vie de l'intervention, et s'appuie sur les grandes dimensions du modèle de l'intervention. Il en découle des études thématiques qui commandent des questions d'évaluation particulières.

Aidée d'un tel cadre, l'équipe d'évaluation peut recevoir les propositions de questions d'évaluation des autorités et des parties prenantes et les situer dans une perspective d'ensemble, pour ensuite répondre à ces demandes de manière fondée et convaincante. L'équipe peut ainsi structurer son plan d'investigation et pondérer les ressources dont elle dispose de manière efficace et efficiente<sup>1</sup>. Elle gagne ainsi en crédibilité face aux autorités responsables et aux parties prenantes.

Dans ce chapitre, nous distinguons d'abord deux grands types de questionnement en évaluation. Nous proposons ensuite les grandes questions d'évaluation découlant des études thématiques du processus d'évaluation. Enfin, un cas de figure servira à illustrer la manière de rédiger les questions d'évaluation.

Nous voulons ainsi tisser un lien étroit et logique entre le modèle de l'intervention, le processus d'évaluation et les questions d'évaluation. Il importe que les questions d'évaluation, qui détermineront en dernière instance l'exercice concret d'évaluation, soient élaborées à partir d'un cadre rigoureux assurant un examen systématique et significatif des interventions publiques.

---

1. Voir le chapitre 11 : *La matrice d'évaluation*.

## Types de questionnements

Si la thématique de chaque étude du processus d'évaluation commande certaines questions d'évaluation, il est une caractéristique incontournable de l'évaluation d'adopter, quelle que soit l'étude thématique, deux grands types de questionnements, distincts et complémentaires : le questionnement descriptif et le questionnement normatif. Pour chaque étude thématique, le questionnement descriptif est essentiel. Selon l'étude cependant, le questionnement normatif peut être soit pertinent, soit prématuré.

### a. Le questionnement descriptif

Le questionnement descriptif de l'évaluation établit un contact factuel et analytique avec l'intervention. Il s'interroge sur la nature, les caractéristiques et les propriétés de l'intervention par :1) l'observation; 2) l'analyse. L'**observation**<sup>2</sup> est l'« acte d'examiner et de noter une chose ou un phénomène par des sens, des instruments ou des procédures ». L'**analyse descriptive** est réalisée « sur la base d'observations visant à représenter un tout à partir de ses éléments ». Par ses questions descriptives, l'évaluation cherche à gagner en connaissances factuelles et en compréhension.

### b. Le questionnement normatif

Le questionnement normatif de l'évaluation vise à traduire la connaissance et la compréhension de l'intervention en valeur sociale. C'est cette interprétation de la réalité, dans ses dimensions humaine et sociale, qui permet l'accomplissement du processus d'évaluation, consacré à estimer la valeur de l'intervention pour la société. Ce questionnement normatif est fondé sur les réponses aux questions descriptives. Mais le questionnement normatif se distingue du questionnement descriptif par son **analyse normative**, définie comme étant l'« analyse basée sur des critères d'évaluation et des standards dans l'intention de porter un jugement ». Les réponses aux questions normatives ajoutent ainsi des constats sur l'adéquation de l'intervention à des attentes ou à des balises établies pour estimer la valeur de l'intervention. Cette vision de « normes » se pose en recul par rapport à l'intervention observée et décrite pour y ajouter la dimension « jugement sur la valeur », c'est-à-dire la dimension « évaluation » proprement dite.

En conséquence, le processus d'évaluation doit combiner intelligemment les questions descriptives et normatives pour connaître, comprendre et juger l'intervention.

La suite de ce chapitre présente les grandes questions descriptives et normatives de l'évaluation.

---

2. La définition du terme observation ainsi que celles d'autres termes d'évaluation non encore mentionnés dans le manuel et marqués ici en gras se trouvent dans le glossaire du chapitre 2 : *La terminologie de l'évaluation*.

## Les grandes questions d'évaluation

Dans le cas souvent espéré, l'équipe d'évaluation profite d'une bonne documentation sur l'intervention provenant des responsables de l'intervention. Elle peut alors « recomposer » sans trop de difficultés le modèle de l'intervention dans ses grandes dimensions à partir de cette documentation. Le processus d'évaluation sert alors à vérifier d'abord la qualité de ce modèle de l'intervention en s'interrogeant sur les choix des méthodes et des techniques de cueillette de données ayant servi à l'élaborer, pour ensuite se consacrer à la portion véritablement évaluative du travail, celle associée au questionnement normatif.

En pratique cependant, l'équipe d'évaluation doit souvent pallier elle-même le manque de renseignements, et elle doit réaliser les analyses descriptives nécessaires à l'élaboration du modèle de l'intervention. Il est alors utile de concevoir ce travail comme une occasion :

1. de reproduire le modèle implicite d'intervention tel que développé par les responsables de l'intervention;
2. de composer un modèle de l'intervention selon une perspective dégagée et objective en respectant les exigences méthodologiques reconnues;
3. d'analyser de manière critique le modèle de l'intervention provenant des responsables de l'intervention.

La partie descriptive pourra alors être considérée comme complétée. Par la suite, l'équipe d'évaluation doit se consacrer à la partie normative du processus d'évaluation.

Bien que les évaluateurs puissent compléter avec profit l'« intelligence » de l'intervention en ajoutant des éléments manquants au modèle d'intervention, ils ne doivent pas confondre leur rôle d'évaluateur avec celui de concepteur ou de maître d'œuvre de l'intervention.

Par exemple, dans le cas d'une intervention dont le référentiel ne présente pas d'objectifs précis, le rôle de proposer des objectifs ne revient pas en première instance à l'équipe d'évaluation. L'absence d'objectifs constitue plutôt un constat du questionnement descriptif devant apparaître comme tel dans le rapport d'évaluation. Le questionnement normatif peut en revanche amener les évaluateurs à recommander la définition d'objectifs dans une version subséquente du référentiel aux autorités responsables de l'intervention. Le rapport d'évaluation peut également suggérer des standards permettant d'établir des attentes réalistes et souhaitables de résultats.

Le tableau 1 présente les grandes questions d'évaluation regroupées par étude thématique. Seules les grandes questions d'évaluation sont présentées. Chaque thématique conduit à développer l'étude en plusieurs questions et sous-questions, chacune pouvant être adaptée au contexte de l'intervention et à celui de l'évaluation.

**Tableau 1** — Les grandes questions d'évaluation regroupées par étude thématique

Étude thématique	Questions d'évaluation
Étude de la raison d'être	Les questions sur la raison d'être de l'intervention : 1. Quelle est la problématique de l'intervention? ( <i>question descriptive</i> ) 2. Quelle est la justification d'intervenir sur la problématique? ( <i>question normative</i> )
Étude de la conception	Les questions sur la conception de l'intervention : 3. Quels sont les intentions, les cibles et les objectifs de l'intervention? ( <i>question descriptive</i> ) 4. Quels sont la nature de l'intervention, la stratégie d'intervention et les instruments d'intervention? ( <i>question descriptive</i> ) 5. Quelles sont les justifications des intentions, des cibles et des objectifs? ( <i>question normative</i> ) 6. Quelles sont les justifications de la nature de l'intervention, de la stratégie d'intervention et des instruments d'intervention? ( <i>question normative</i> )
Étude de la mise en œuvre Étude de la conformité	Les questions sur la mise en œuvre de l'intervention : 7. Quel est le plan de mise en œuvre, quels sont les cibles opérationnelles et les objectifs opérationnels? ( <i>question descriptive</i> ) 8. Quels sont les intrants, l'activité de production et les extrants? ( <i>question descriptive</i> ) 9. Quelle est la conformité de la mise en œuvre avec le plan de mise en œuvre? ( <i>question normative</i> )
Étude des coûts	10. Quels sont les coûts de production et les coûts unitaires de production? ( <i>question descriptive</i> )
Étude de la performance opérationnelle	11. Quelles sont l'efficacité opérationnelle, l'efficience opérationnelle et l'équité opérationnelle de l'intervention? ( <i>question normative</i> )
Étude des effets	Les questions sur les effets de l'intervention : 12. Quels sont les effets de l'intervention? ( <i>question descriptive</i> )
Étude de la valeur de l'intervention	Les questions sur la valeur de l'intervention : 13. Quelle est l'efficacité globale de l'intervention? ( <i>question normative</i> ) 14. Quelle est l'efficience globale de l'intervention? ( <i>question normative</i> ) 15. Quelle est l'équité globale de l'intervention? ( <i>question normative</i> ) 16. Quelle est la valeur de l'intervention? ( <i>question normative</i> )

Une numérotation couvrant l'ensemble des questions du tableau 1 rappelle la nature séquentielle du processus d'évaluation, des études thématiques et des questions d'évaluation. Le statut descriptif ou normatif du questionnement est indiqué pour chaque question. Il est notable que chaque étude thématique comporte au moins une question descriptive, mais pas nécessairement une question normative. Le cas de l'étude des effets est particulièrement illustratif : la seule grande question qu'elle comporte est de nature descriptive. La tâche de documenter les effets d'une intervention est fort exigeante et mérite à la fois une distanciation objective et une logistique importante. L'interprétation des effets d'une intervention occupera cependant l'entièreté de l'étude de la valeur de l'intervention grâce à ses quatre questions normatives.

### a. L'étude de la raison d'être

Cette première étude thématique est consacrée à la raison d'être de l'intervention, soit « l'ensemble des motifs qui légitime une intervention ». Elle amène l'équipe d'évaluation à aller aux sources de la volonté d'agir pour en connaître et en comprendre les fondements et les justifications. L'étude de la raison d'être exige donc : 1) de poser des questions et des sous-questions de nature descriptive pour connaître et comprendre la problématique suscitant une intervention; 2) de soulever d'autres questions de nature normative permettant de connaître et de valider la justification de l'intervention.

Le questionnement sur la raison d'être peut se développer sous la forme suivante :

1. Quelle est la problématique de l'intervention? (*question descriptive*)
2. Quelle est la justification d'intervenir sur la problématique? (*question normative*)

Il est certes souhaitable que l'autorité responsable de l'intervention ait pu exprimer publiquement et clairement la raison d'être de l'intervention avant de la concevoir et de la mettre en œuvre. Mais il ne faut pas le présumer. L'équipe d'évaluation doit documenter les prises de position explicites et officielles à l'égard de la raison d'être, mais il doit être envisagé de recomposer le processus de décision autour de la raison d'intervenir si la documentation est insuffisante.

Mais plus encore, l'équipe d'évaluation doit examiner les motifs invoqués pour justifier l'intervention. À ce chapitre, l'évaluation autorise l'analyse critique des positions officielles de l'autorité, ce que ne permet pas en général l'audit ou la vérification.

Ce rôle plus critique de l'évaluation tient à sa définition même. Estimer la valeur d'une intervention pour la société exige un examen approfondi de la problématique qui a servi de justification à l'intervention. S'il s'avérait que la problématique d'intervention soit mal établie ou mal fondée, ou que la justification d'intervenir ne soit pas défendable, il est du rôle de l'évaluation de révéler cette faiblesse de l'intervention. Dans un tel cas, nier ce rôle reviendrait à procéder à un examen biaisé, et en plus inefficace, de tout le processus de conception et de mise en œuvre d'une intervention.

L'étude de la raison d'être exige donc de déployer les méthodes et techniques appropriées pour répondre à la question thématique : quelle est la raison d'être de l'intervention ?

### b. L'étude de la conception

La deuxième étude thématique est consacrée à la conception de l'intervention, soit le « processus visant à définir les intentions et la nature de l'intervention ». Elle amène les évaluatrices et les évaluateurs à connaître et à comprendre la manière dont l'autorité a transformé sa compréhension de la problématique en une stratégie et en une série d'actions ordonnées, planifiées et crédibles pour résoudre les difficultés que pose la problématique. L'étude de la conception passe d'abord par un questionnement descriptif des intentions d'agir et de la nature de l'intervention. Mais elle se poursuit par une analyse rigoureuse de la manière dont les intentions ont été formulées et de la façon dont les choix de l'autorité quant à la nature de l'intervention, à la **stratégie d'intervention** et aux instruments d'intervention, ont été faits et justifiés.

L'étude de la conception peut se développer autour des questions suivantes :

3. Quels sont les intentions, les **cibles** et les **objectifs de l'intervention**? (*question descriptive*)
4. Quels sont la nature de l'intervention, la stratégie d'intervention et les instruments d'intervention? (*question descriptive*)
5. Quelles sont les justifications des intentions, des cibles et des objectifs? (*question normative*)
6. Quelles sont les justifications de la nature de l'intervention, de la stratégie d'intervention et des instruments d'intervention? (*question normative*)

Le questionnement descriptif de l'étude de la conception enrichit la modélisation de l'intervention au stade de la réflexion qui a précédé la mise en œuvre. Le questionnement normatif exige quant à lui de connaître le raisonnement et les justifications du passage de la problématique d'intervention à la fixation de cibles et d'objectifs, et de juger de leur valeur. Le questionnement normatif appelle également à comprendre le passage entre des intentions claires et justifiées et le choix d'une stratégie d'intervention et des instruments d'intervention.

### c. L'étude de la mise en œuvre

La troisième étude thématique est consacrée à la mise en œuvre de l'intervention, soit le « processus visant à concrétiser l'intervention ». Cette étude amène l'équipe d'évaluation à connaître et à comprendre ce processus qui fait passer l'autorité de la parole aux actes! C'est l'univers des opérations qui suppose un **plan de mise en œuvre**, des **cibles opérationnelles** et des **objectifs opérationnels**, des **intrants**, des **coûts de production** et des **coûts unitaires de production**, une **activité de production** et des **extrants**. L'étude de mise en œuvre doit faire le point et doit dresser un bilan sur l'état de la situation grâce à une étude de la **performance opérationnelle**.

Par l'étude de la conformité, l'équipe d'évaluation cherche des réponses aux trois questions suivantes :

7. Quel est le plan de mise en œuvre et quels sont les objectifs opérationnels? (*question descriptive*)
8. Quels sont les intrants, l'activité de production et les extrants? (*question descriptive*)
9. Quelle est la conformité de la mise en œuvre avec le plan de mise en œuvre? (*question normative*)

Les deux premières questions de l'étude de la conformité sont de nature descriptive. La troisième question, de nature normative, statue d'abord sur le degré de conformité de la mise en œuvre avec le plan de mise en œuvre. Elle doit ensuite expliquer les écarts avec le plan de mise en œuvre, car il demeure plausible que la mise en œuvre déroge du plan de mise en œuvre pour des motifs valables.

L'estimation des coûts de production et des coûts unitaires de production (l'équivalent du prix de revient) permet, entre autres choses, l'appréciation ultérieure de la performance opérationnelle.

10. Quels sont les coûts de production et les coûts unitaires de production? (*question descriptive*)

Enfin, l'étude de la performance opérationnelle apporte un point de vue normatif à l'étude de la mise en œuvre en s'interrogeant sur le respect de critères de performance en termes **d'efficacité opérationnelle**, **d'efficience opérationnelle** et **d'équité opérationnelle**.

11. Quelles sont l'efficacité opérationnelle, l'efficience opérationnelle et l'équité opérationnelle de l'intervention? (*question normative*)



#### d. L'étude des effets

La quatrième étude thématique est consacrée aux effets de l'intervention, un effet étant un « changement qui est produit par une intervention ». Cette étude se concentre sur l'observation, la mesure et l'analyse de la contribution de l'intervention aux changements observés dans la société. Les difficultés pratiques de distinguer, à travers l'évolution des situations sociales et de la problématique d'intervention, les changements dus spécifiquement à l'intervention, font de cette étude une des plus exigeantes. Elle apporte cependant des évidences et des démonstrations essentielles à l'évaluation de l'intervention qu'aucune étude précédente n'est en mesure d'offrir. L'intervention existant pour résoudre une problématique, l'étude des effets vise à observer tout changement dû à l'intervention de manière à statuer ultérieurement sur sa capacité à résoudre la problématique.

##### 12. Quels sont les effets de l'intervention ? (*question descriptive*)

Par cette thématique descriptive, l'étude des effets n'a pas à juger quoi que ce soit : elle ne pose donc pas de question normative. Décrire et expliquer les effets de l'intervention nécessitent expertise, temps et ressources. Elle fournira des informations cruciales à la cinquième et dernière étude thématique qui, elle, a une finalité essentiellement normative.

Il est cependant intéressant et souvent pertinent de distinguer plusieurs types d'effets, car chaque type d'effet porte une signification particulière et contribue à sa manière aux constats et aux recommandations de l'évaluation :

- i. Les **effets sur cible directe**, les **effets sur cible intermédiaire** et les **effets sur cible ultime**
- ii. Les **effets hors cible**
- iii. Les **effets prévus** et les **effets non prévus**
- iv. Les **effets intentionnels** et les **effets non intentionnels**

#### e. L'étude de la valeur de l'intervention

La cinquième et dernière étude thématique est consacrée à la valeur de l'intervention, soit « la valeur d'une intervention pour la société ». Cette étude voit à faire une synthèse des observations et des constats descriptifs et normatifs obtenus tout au long du processus d'évaluation.

##### 13. Quelle est l'**efficacité globale** de l'intervention ? (*question normative*)

##### 14. Quelle est l'**efficience globale** de l'intervention ? (*question normative*)

##### 15. Quelle est l'**équité globale** de l'intervention ? (*question normative*)

##### 16. Quelle est la valeur de l'intervention ? (*question normative*)

En particulier, l'étude de la valeur de l'intervention doit juger de l'efficacité globale<sup>3</sup> de l'intervention en comparant les effets de l'intervention à ses objectifs et à la résolution de la problématique. Elle doit également juger de l'efficience globale en mettant en rapport les effets positifs de l'intervention aux divers coûts et sacrifices qu'elle a imposés à la société. Elle doit aussi faire une lecture de l'équité globale de l'intervention, en ce sens que les avantages et les inconvénients de l'intervention se partagent souvent de manière inégale dans la société. Ainsi renseignée, l'équipe d'évaluation est à même de conclure sur la valeur de l'intervention, en rappelant toutefois les limites inévitables du travail d'évaluation.

3. Globale étant entendu au sens « d'ensemble » plutôt qu'au sens anglais de « planétaire ».

## La considération de questions particulières

Les questions d'évaluation proposées ici se veulent d'une certaine universalité, car elles proviennent de la nature même d'une intervention sur la société. Mais chaque société, chaque culture et chaque époque affectent à leur manière le questionnement sur l'intervention. Aussi faut-il se demander de quelle façon le processus d'évaluation proposé, les études thématiques et les questions d'évaluation peuvent refléter et peuvent s'adapter aux particularités de chaque lieu et circonstance. Indéniablement, toute proposition de processus d'évaluation et de questions d'évaluation doit offrir la flexibilité pour s'adapter à l'évolution de la pensée en évaluation et à l'évolution des sociétés.

Par exemple, il est difficile à notre époque d'évaluer une intervention sans l'examiner, par exemple, sous l'angle du développement durable ou sous l'angle des genres. Comment, par exemple, ces deux questions d'une pertinence indéniable aujourd'hui, et comment d'autres questions qui s'imposeront dans un avenir proche ou éloigné, peuvent-elles s'introduire de manière cohérente et fructueuse dans le processus d'évaluation proposé ici ?

Il est bien sûr possible d'ajouter à la liste de questions du tableau 1 une liste supplémentaire de questions : une sorte d'adaptation additive. Mais le processus d'évaluation, tel que proposé, et le questionnement ainsi ajouté, perdraient en qualité en n'intégrant qu'artificiellement ces questions « supplémentaires » aux « premières » questions qui profitent déjà d'une prise en compte tout au long du processus d'évaluation.

Heureusement, il y a mieux. Les critères d'efficacité, d'efficience et d'équité utilisés ici, tant au niveau opérationnel qu'au niveau global, peuvent être entendus comme des supercritères pouvant englober un ensemble plus vaste de critères ou de préoccupations caractéristiques d'une époque ou d'un lieu donné. Or, tout le processus d'évaluation prépare l'examen de l'intervention sous l'angle de ces supercritères. Il y a donc avantage à situer de nouvelles préoccupations ou de nouveaux critères à l'intérieur des supercritères pour une considération complète et efficace, lorsque, bien sûr, cela est possible.

Prenons le cas du questionnement relatif au développement durable. Il peut être rigoureusement interprété comme une préoccupation d'**efficience**<sup>4</sup> selon un horizon temporel « élargi » ou « durable », qui incorpore à la fois les besoins des générations actuelles, mais aussi ceux des générations futures (comme le fait la technique d'analyse coût-avantage). En effet, le critère d'efficience globale correspond à la « performance relative au rapport entre les **avantages sociaux** et les **coûts sociaux** ». Rien dans ce concept n'empêche la comparaison des avantages sociaux aux coûts sociaux sur un horizon temporel étendu. Au contraire même, on peut se demander pourquoi nous évaluerions l'efficience globale d'une intervention sur un horizon temporel limité, soit celui des générations actuelles, sinon par une sorte de myopie temporelle, une pensée à courte vue<sup>5</sup>. Aussi le questionnement sur la durabilité des interventions publiques s'incorpore-t-il naturellement dans l'examen de l'efficience globale de l'intervention, à la condition cependant d'être attentif et conséquent quand vient le moment de traiter de l'horizon temporel dans l'appréciation des avantages et des coûts sociaux.

En situant logiquement et naturellement le concept de développement durable à l'intérieur du critère d'efficience globale, ce questionnement profitera de la cohérence du processus d'évaluation, ce qui facilitera sa prise en compte systématique et entière dans le processus d'évaluation.

4. Le même raisonnement sur le développement durable pourrait être fait sous l'angle de l'équité globale.

5. Voir par exemple Pierre Rosanvallon, Sortir de la myopie des démocraties, *Le Monde*, 7 décembre 2009.

Il en va de même par exemple pour la préoccupation des genres. Bien que cette préoccupation puisse prendre plusieurs formes, considérant également que la notion même de genre est en pleine redéfinition, il est avantageux, en termes de cohérence et aussi de pertinence pour l'évaluation, de considérer cette préoccupation comme une préoccupation d'équité. L'**équité** est un concept général qui correspond en évaluation à la « performance relative à la capacité d'une action à être juste ».

Ce concept appliqué à la mise en œuvre d'une intervention, l'équité opérationnelle, devient la « performance relative aux justes conditions des personnes impliquées dans la mise en œuvre ». Ainsi, il contient naturellement la préoccupation des genres qui peut être avantageusement explicitée dans les sous-questions sur la performance opérationnelle. Elle peut être traduite ainsi : « Est-ce que la mise en œuvre de l'intervention s'est montrée performante quant à l'équité des conditions des personnes impliquées sur la base des genres ? »

L'équité globale d'une intervention, quant à elle, correspond à la « performance relative à une juste distribution des avantages sociaux et des coûts sociaux ». Ce même concept appliqué à la question des genres peut se traduire ainsi : « Est-ce que l'intervention s'est montrée performante quant à la juste distribution des avantages sociaux et des coûts sociaux sur la base des genres ? ».

Ainsi, la question des genres en évaluation gagne aussi à être incorporée naturellement et entièrement dans le processus d'évaluation proposé ici. Les études thématiques et les questions d'évaluation préalables préparent au mieux le processus d'évaluation et évitent que ces questions soient traitées (ou même maltraitées) à la toute fin d'un processus sans bénéficier de tout l'appui du processus évaluatif.

On peut aisément étendre ce raisonnement à toutes les préoccupations sur l'équité envers les groupes sociaux qui pourraient être défavorisés par les résultats des interventions publiques (ex. : race, ethnie, religion, âge, etc.).

En conclusion, les évaluateurs auront au fil du temps à contribuer à l'amélioration et à la précision du processus d'évaluation et des grandes questions d'évaluation proposés ici. Ce faisant, ils auront avantage à conserver la plus grande cohérence possible au processus d'évaluation pour lui conserver son efficacité et sa pertinence tout en facilitant ainsi la réalisation, la compréhension et l'acceptation de leurs travaux d'évaluation.

## Une illustration de rédaction des questions d'évaluation

Nous terminons ce chapitre en illustrant la rédaction des questions d'évaluation par le cas fictif d'une intervention de lutte contre la pauvreté :

« *Une intervention en matière de lutte contre la pauvreté : cas fictif* » (Vénétia Sauvain, École nationale d'administration publique, 2020, adapté de l'Initiative nationale pour le développement humain du Royaume du Maroc).

Contexte du cas :

*La prise de conscience d'une problématique de pauvreté, s'exprimant souvent sous la forme de précarité, d'exclusion sociale, de chômage, de délinquance et de misère, et touchant particulièrement les femmes et les jeunes, amène l'État à intervenir pour améliorer significativement la situation. Les niveaux ministériel, régional et local sont amenés à agir avec l'intention de diminuer l'ampleur de la pauvreté dans le pays. L'intervention vise à financer des projets d'activité génératrice de revenus (AGR) et de microprojets coopératifs (MPC). Un système de suivi-évaluation de l'intervention est mis en place avec la collaboration des différents paliers d'intervention.*

N. B. : Le traitement du cas adopte ici une structure établie selon le processus d'évaluation proposé et les grandes questions d'évaluation pour établir le modèle de l'intervention. La colonne de gauche reprend, sous forme d'éléments du modèle de l'intervention, une partie des questions proposées ici et répond de façon synthétique dans la colonne de droite.

Cette illustration veut montrer comment on peut appliquer à un cas particulier les grandes questions d'évaluation. Les exemples de questions proposés au tableau 2 peuvent ne pas convenir à des situations régionales particulières. Ils veulent surtout illustrer la flexibilité des grandes questions à s'adapter à différents contextes, et à inviter les équipes d'évaluation à imaginer les questions et sous-questions pertinentes dans leur situation et qui conduiront l'exercice d'évaluation.

Le tableau 2 énonce l'ensemble des questions d'évaluation proposées au tableau 1 de ce chapitre en prenant soin de les adapter le plus possible au cas fictif. La numérotation des questions du tableau 2 suit l'ordre des grandes questions d'évaluation du tableau 1.

**Tableau 2 — Illustration d'un questionnement évaluatif au cas fictif : le Programme de lutte contre la pauvreté urbaine et rurale (PGLPUR)**

Étude thématique	Hypothèses de questions et de sous-questions d'évaluation
Étude de la raison d'être	<p>Les questions sur la raison d'être du PGLPUR :</p> <ol style="list-style-type: none"> <li>1. Quelle est la problématique du PGLPUR? Quel est le problème principal ou quels sont les problèmes principaux, leur nature, leurs causes et leurs conséquences? Quels sont les angles ou catégories sociales particulièrement affectés par la problématique en question? Quel est l'horizon temporel considéré par la problématique? Quel est le modèle de la problématique qui s'en dégage?</li> <li>2. Quelle est la justification d'intervenir sur cette problématique de pauvreté? Pourquoi le marché de l'emploi ne peut-il pas de lui-même régler le problème de la pauvreté en zone urbaine? En zone rurale? Quel est le rôle de l'État et des divers paliers quant à cette problématique de pauvreté? L'État doit-il assumer l'entière responsabilité et des actions correctives, ou doit-il les partager en partie avec le marché, avec les autres paliers? Pourquoi?</li> </ol>
Étude de la conception	<p>Les questions sur la conception du PGLPUR :</p> <ol style="list-style-type: none"> <li>3. Quelles sont les intentions du gouvernement, sa volonté d'agir dans cette problématique? Quelles sont les cibles directes (immédiates), les cibles intermédiaires (à moyen terme) et les cibles ultimes du programme? Ces cibles reflètent-elles bien les catégories sociales identifiées comme principalement affectées par la problématique? Quels sont les objectifs concernant chacune des cibles identifiées précédemment?</li> <li>4. Quelle est la nature de l'intervention retenue? Quelle est la stratégie d'intervention et quels sont les instruments d'intervention?</li> <li>5. Quelles sont les justifications des intentions annoncées, des cibles et des objectifs?</li> <li>6. Quelles sont les justifications concernant le choix de la nature de l'intervention, de la stratégie d'intervention et des instruments d'intervention? Quelle est la théorie d'intervention qui se dégage de la logique de l'intervention et des moyens retenus, et quels sont les effets prévus de la mise en œuvre du PGLPUR?</li> </ol>

Étude thématique	Hypothèses de questions et de sous-questions d'évaluation
<p>Étude de la mise en œuvre Étude de la conformité</p>	<p>Les questions sur la mise en œuvre du PGLPUR :</p> <p>7. Quel est le plan de mise en œuvre du PGLPUR tant au niveau national qu'au niveau de chaque palier gouvernemental impliqué? Quels sont les objectifs opérationnels exprimés pour chaque intervenant du PGLPUR?</p> <p>8. Quels sont les intrants prévus et accessibles pour la mise en œuvre du PGLPUR? Quelle est l'activité de production et quels sont les extrants? Quelle est donc la chaîne de production prévue et réelle?</p> <p>9. Quelle est la conformité de la mise en œuvre avec le plan de mise en œuvre? Quels sont les écarts? Comment ces écarts sont-ils expliqués?</p>
<p>Étude des coûts</p>	<p>10. Quels sont les coûts totaux et les coûts unitaires de production du PGLPUR?</p>
<p>Étude de la performance opérationnelle</p>	<p>11. Quelles sont l'efficacité opérationnelle (atteinte des objectifs opérationnels), l'efficience opérationnelle et l'équité opérationnelle (genres, régions, catégories sociales...) du PGLPUR?</p>
<p>Étude des effets</p>	<p>Les questions sur les effets du PGLPUR :</p> <p>12. Quels sont les effets du PGLPUR? Quels sont les effets sur les cibles directes, intermédiaires et ultimes? Quels sont les effets hors cible? Quels sont les effets non intentionnels? Quels sont les effets sur les changements climatiques, sur l'adaptation aux changements climatiques?</p>
<p>Étude de la valeur de l'intervention</p>	<p>Les questions sur la valeur de l'intervention :</p> <p>13. Quelle est l'efficacité globale de l'intervention (atteinte des objectifs du PGLPUR)?</p> <p>14. Quelle est l'efficience globale de l'intervention (rapport entre les avantages et les coûts pour la société) en considérant un horizon temporel impliquant les générations futures?</p> <p>15. Quelle est l'équité globale de l'intervention, en considérant la distribution des avantages et des coûts du PGLPUR entre les genres, entre les pauvres et les mieux pourvus, et d'autres catégories sociales?</p> <p>16. Quelle est la valeur du PGLPUR, tout bien considéré pour la population du pays?</p>

## Messages clés

1. Les grandes questions d'évaluation découlent des études thématiques du processus d'évaluation.
2. Puisque les études thématiques sont imbriquées dans une séquence précise, les questions se positionnent les unes par rapport aux autres de manière séquentielle.
3. Deux types de questions composent le questionnement évaluatif : les questions descriptives et les questions normatives.
4. Le spectre des questions devrait couvrir à la fois ce qui précède l'intervention, l'intervention elle-même et ses conséquences.
5. Pour des raisons pratiques, les questions peuvent se regrouper en quelques études pouvant être réalisées à priori, in itinéré ou à postériori.
6. Des questions particulières, relatives à des préoccupations contextuelles, sociales ou historiques, peuvent exiger une grande attention. Elles méritent alors d'être intégrées aux grandes questions d'évaluation d'une manière cohérente pour profiter ainsi d'un traitement en profondeur dès le déclenchement de l'exercice d'évaluation.
7. Par exemple, les enjeux de développement durable ou de genre peuvent être intégrés comme considérations particulières d'efficacité et d'équité, tout en conservant leur identité et leur visibilité dans l'exercice d'évaluation.

## Exercice de réflexion

Le mandat qui est confié à votre équipe consiste à évaluer une initiative majeure de développement visant l'approvisionnement en eau potable d'une partie significative de la population du pays.

N. B. : les mises en situation sont indépendantes les unes des autres.

1. La problématique à laquelle l'initiative est consacrée serait, selon le mandant, très bien documentée par des experts. Vous obtenez copie de ces rapports d'experts. Étant donné cet avis du mandant et les opinions des experts sur la problématique, comment voyez-vous alors votre rôle d'évaluateur dans l'étude de la raison d'être de l'initiative ?
2. Un haut responsable de l'initiative vous explique que l'évaluation ne conduira à rien puisque, lors de la conception de l'initiative, aucun objectif précis n'a été établi. L'atteinte des objectifs (efficacité) serait donc impossible à estimer et, en conséquence, l'évaluation serait impertinente. Que pensez-vous de ce raisonnement ?
3. Vous constatez que le choix de la stratégie d'intervention et des instruments d'intervention n'est pas documenté, et les entrevues que vous menez ne vous apprennent rien sur les autres options qui auraient pu être considérées. Quelle décision prendrez-vous à cette étape du processus d'évaluation ?
4. Votre étude de la conformité révèle que la mise en œuvre ne s'est pas effectuée comme prévu. Les hauts responsables de l'initiative se montrent irrités de la situation et sont prêts à se montrer sévères face au maître d'œuvre. Quelle suite donnerez-vous à l'étude de la mise en œuvre ?

5. Lors de l'étude des effets, il apparaît que les effets sont nettement inférieurs à ceux escomptés. Les bailleurs de fonds de l'évaluation vous proposent d'arrêter le processus d'évaluation, étant donné qu'ils ont obtenu leurs réponses. Quelle sera votre réplique à cette proposition ?
6. L'initiative aurait, selon votre étude de la valeur de l'intervention, atteint ses objectifs, et en conséquence montrerait un degré élevé d'efficacité globale. Les hauts responsables de l'initiative s'en félicitent et suggèrent là encore d'arrêter le processus d'évaluation. Pourquoi continuer ? Quelle sera alors votre réponse ?
7. L'initiative montre un degré élevé d'efficacité globale. Au total, la communauté obtient par cette intervention plus d'avantages qu'elle ne subit de coûts. Que nous apprendra de plus la connaissance de l'équité globale ?
8. Advenant que l'équité globale de l'initiative soit en deçà des intentions de l'autorité, mais que l'efficacité de l'initiative soit excellente, comment établirez-vous, en conclusion de votre rapport, la valeur de l'intervention ?





# Partie II

## Phase méthodologique



# La matrice d'évaluation

*Richard Marceau*



## Introduction

En évaluation comme en d'autres choses, il est un temps pour la réflexion ... et un temps pour l'action. La matrice d'évaluation, objet de ce chapitre, marque la transition entre la réflexion et l'action : elle reprend, d'une part, les questions d'évaluation (présentées au chapitre 10 : *Les questions d'évaluation*) et, d'autre part, propose pour chacune une démarche précise pour obtenir des réponses. La matrice d'évaluation tire son nom de sa structure bidimensionnelle : un axe vertical comportant les questions d'évaluation et un axe horizontal déployant la démarche de recherche de réponses. Ces deux axes donnent lieu à des cellules contenant les propositions concrètes pour la suite du processus d'évaluation. Sa fonction est de rassembler et de coordonner la logistique de l'évaluation en un seul tableau pratique qui guide la concrétisation du processus d'évaluation (voir chapitre 9 : *Le processus d'évaluation*).

La matrice d'évaluation n'est pas qu'un simple marqueur temporel dans la série d'étapes du processus d'évaluation : elle correspond au moment où les décisions logistiques doivent être prises. C'est l'occasion d'introduire les possibilités et les contraintes méthodologiques et contextuelles dans le plan d'évaluation.

Par l'exercice de « réalisme » qu'impose l'élaboration de la matrice d'évaluation, les possibilités et les contraintes affecteront la qualité et la couverture des réponses aux questions. Il s'ensuit que la pertinence, la profondeur et la qualité de l'évaluation seront tributaires de l'ajustement logistique aux questions, tel qu'exprimé dans la matrice d'évaluation. Le professionnalisme avec lequel cet ajustement logistique sera réalisé permettra, ou non, de maximiser la valeur et, éventuellement, l'utilisation de l'évaluation.

Les compétences méthodologiques et techniques des évaluateurs et leur connaissance des coûts de réalisation seront pleinement interpellées dans cet exercice pour faire en sorte que les choix logistiques permettent à l'évaluation de répondre au mieux aux attentes des autorités et des parties prenantes.

Au bout de cette construction matricielle se visualisera la transition entre l'abstrait et le concret de l'évaluation. L'assemblage des solutions logistiques à chaque question posée concrétisera la démarche évaluative.

## Une démarche de type « recherche appliquée »

La décomposition du processus d'évaluation en études thématiques et en grandes questions (comme nous l'avons fait au chapitre 9 : *Le processus d'évaluation*, et au chapitre 10 : *Les questions d'évaluation*) définit en partie la structure de la démarche évaluative.

D'abord, les questions évaluatives sont de deux ordres : les questions descriptives et les questions normatives. À cet égard, les questions descriptives doivent obtenir des réponses avant que les questions normatives soient posées.

Ensuite, les questions d'évaluation appellent des réponses, mais pas n'importe quelles réponses : des réponses précises, développées, crédibles et vérifiables. Or, en termes de qualité de réponses, il existe plusieurs standards. Le standard peut être celui de l'amateur – qui retient des éléments de réponse valables –, ou au contraire fallacieux, sans mécanisme rigoureux de tri, celui du journaliste – procédant par recoupement de récits et de documents –, ou celui de la démarche scientifique, et bien d'autres. Chacun de ces standards implique des compétences et des ressources particulières.

### *Standards de qualité en évaluation*

Pourquoi ne pas choisir systématiquement les méthodes et outils du plus haut niveau de qualité ? Parce qu'en général, ils coûtent plus cher. Or, les interventions publiques elles-mêmes coûtent cher à la société. L'évaluation n'est qu'un poste de dépenses parmi d'autres, et il est clair que ce qui est épargné en évaluation pourrait servir à augmenter les ressources consacrées à l'intervention. Un arbitrage doit être réalisé entre le budget consacré à l'intervention et celui consacré à son évaluation. Les choix méthodologiques et techniques de l'évaluation en seront ainsi affectés.

Il semblerait à priori logique de financer les évaluations en proportion des coûts des interventions pour que leur standard de qualité soit alors du plus haut niveau possible. L'arbitrage budgétaire entre l'intervention et son évaluation serait ainsi fixé à l'avance. Mais il n'en va pas ainsi. Certaines interventions coûteuses peuvent s'évaluer par des moyens qui profitent de grandes économies d'échelle. D'autres interventions, plus modestes, exigent des moyens particuliers et coûteux. Comment alors, devant ces difficultés, peut-on définir des standards de qualité en évaluation ? Comment les évaluateurs peuvent-ils à la fois respecter un budget donné et des standards de qualité élevés ?

L'analogie avec la pratique médicale peut être appropriée ici.

Règle générale, le médecin ne contrôle pas lui-même les ressources disponibles pour son acte médical. On exige de lui la plus haute compétence médicale, mais son acte sera établi en fonction des ressources dont il dispose et du contexte de son intervention. Il classera les moyens d'intervention médicale en les hiérarchisant selon leur qualité et proposera le moyen accessible le plus performant. Il fonctionnera avec les plus hauts standards, mais il interviendra en tenant compte des contraintes qui s'exercent sur ses choix. Il ne sacrifie pas les standards de la médecine : il effectue des choix de moyens réalistes sur la base de hauts standards. Il est en mesure d'expliquer et de défendre ses choix. Il peut alors justifier sa décision sur la base de standards de qualité, mais n'assumera pas lui-même le poids des contraintes qui lui sont imposées dans sa pratique.

Il en va de même en évaluation. Les évaluatrices et évaluateurs doivent connaître les moyens de réponse aux questions évaluatives quant à leurs capacités et à leurs exigences en ressources. Ils doivent savoir proposer les moyens appropriés grâce à une hiérarchisation de la qualité des moyens et une décision fondée sur leurs implications en ressources. Ils ne doivent pas sacrifier leurs standards de

qualité, mais plutôt les utiliser dans un contexte de contraintes pour que des choix méthodologiques et techniques qui correspondent au cadre financier propre à l'évaluation en cours soient judicieusement effectués. Ils n'ont pas à assumer eux-mêmes la responsabilité des contraintes qui pèsent sur l'évaluation, mais doivent en tenir compte dans leurs propositions.

Or, les questions d'évaluation – par exemple celles portant sur le choix des instruments d'intervention et sur leurs effets – interpellent des moyens qui sont typiques de ceux des sciences sociales en général, et parfois aussi de ceux des sciences naturelles : théorie des instruments d'intervention, observations sur le terrain, mesures quantitatives et qualitatives de santé publique, questionnaires, sondages, protocoles d'entrevue, etc. L'évaluation est d'abord un champ de connaissances avant d'être un champ de pratique. La pratique de l'évaluation doit se nourrir du champ de connaissances qui lui fournit la boîte à outils dont elle a besoin pour répondre aux questions évaluatives.

À l'image de la pratique médicale, la pratique évaluative doit adopter une démarche rigoureuse certes, mais, plus précisément, une démarche de type « recherche appliquée ». La recherche appliquée se nourrit de manière générale de la recherche fondamentale et de méthodologies déductives et inductives. Elle est adaptée aux différents contextes d'application. De même, l'évaluation doit se nourrir des plus hauts standards de qualité méthodologiques et techniques, mais doit effectuer des choix particuliers qui sont en accord avec les contraintes de réalisation.

Ainsi, les choix sur les moyens de réponse à chaque question évaluative doivent être effectués sur la base d'une hiérarchie fondée sur les plus hauts standards méthodologiques et techniques, ceux de la démarche scientifique, et sur la base des contraintes de réalisation, comme le fait la démarche de type « recherche appliquée ». En conséquence, il n'existe pas de méthode ou de technique passe-partout qui convienne en tout temps et en toutes circonstances. C'est le contexte qui permet d'identifier les méthodes et outils optimaux dans la boîte à outils de la recherche appliquée et de faire les choix conséquents. Les choix méthodologiques et techniques figurant dans la matrice d'évaluation se doivent d'être contingents au contexte de réalisation de l'évaluation. Les évaluatrices et évaluateurs doivent en toutes circonstances examiner les possibilités méthodologiques et techniques pour proposer les solutions les plus appropriées. Cette « capacité de faire » est leur valeur ajoutée.

#### *Une démarche de recherche appliquée*

La démarche de type « recherche appliquée » en évaluation peut prendre la forme suivante :

1. Une **question** est posée.
2. Une **hypothèse** de réponse, ancrée ou non dans une théorie (théorie du problème, théorie de l'intervention...), prend la forme d'un état de situation qui appelle des observations et une vérification d'hypothèse. Ce peut être à propos d'une opinion, d'une attitude, d'un comportement, d'un événement, d'une situation environnementale, d'un résultat, ou à propos de causes et de conséquences d'une situation. À défaut d'hypothèses, une recherche d'hypothèses de réponse à la question se bâtit au fur et à mesure des observations faites sur le terrain.
3. Une ou des **variables** exprimeront ou expliqueront l'état de situation.
4. Une **méthode** est choisie pour vérifier la validité de l'hypothèse de réponse.
5. Des **indicateurs** sont choisis pour l'observation et la mesure de l'état des variables retenues ou de toute autre variable jugée d'importance par la méthode de vérification. Le choix de la méthode et celui des indicateurs sont souvent tributaires l'un de l'autre.

6. Des **outils de collecte de données** sont sélectionnés pour l'observation et la mesure des indicateurs.
7. Des **analyses de données** selon la méthode retenue sont effectuées pour formuler la réponse recherchée à la question initiale.

En cela, la matrice d'évaluation est typique d'une démarche de recherche que l'on retrouve en sciences sociales ou en sciences naturelles (voir tableau 1).

**Tableau 1** — Structure typique d'une démarche de recherche

Démarche de recherche	Théorie/hypothèse	Variables	Méthode	Indicateurs	Outils	Analyses
Questions de recherche						
Question n°1						
Question n°2						
Question n°3						
Question n°4						
Question n°5						

Bien que cette recherche « typique » de réponse aux questions adopte une forme déductive, elle n'exclut pas une recherche de type inductif.

Par exemple, à la question « quels sont les effets de l'intervention ? », une démarche évaluative de type déductif conduira à énoncer une ou des hypothèses sur l'ampleur des effets prévus. Des variables, des méthodes de vérification d'hypothèses, des indicateurs concernant les variables d'effets potentiels et des techniques de collecte de données seront choisis pour permettre les analyses de données vérifiant la ou les hypothèses énoncées sur les effets de l'intervention. Cependant, certaines démarches évaluatives de type inductif conduiront plutôt à élaborer les hypothèses, à choisir les méthodes, les variables et les indicateurs en interaction avec les observations terrain pour ensuite effectuer les analyses de données et la formulation d'une réponse à la question d'évaluation.

La matrice d'évaluation, dans sa forme la plus répandue dans la pratique d'évaluation, retient celle de l'approche déductive, qu'il faut savoir adapter à des approches plus inductives, au besoin.

## Une suite de recherches imbriquées

Le tableau 1 du chapitre 10 : *Les questions d'évaluation* développe 16 grandes questions d'évaluation composant les cinq études thématiques du processus d'évaluation. Développer la structure d'une matrice d'évaluation complète revient concrètement à lister, en première colonne de la matrice, les 16 questions d'évaluation ou une partie d'entre elles, et inscrire en première ligne les six étapes de la démarche évaluative de type « recherche appliquée ».

À partir de cette structure matricielle, le travail de planification de l'équipe d'évaluation consiste à déterminer, pour chaque question d'évaluation, les diverses hypothèses de réponse, les méthodes, outils et indicateurs qui permettent d'apporter une réponse (voir tableau 2). Aussi, la matrice

d'évaluation peut être vue comme une série de recherches appliquées comportant autant de recherches distinctes que de questions d'évaluation. On mesure alors toute l'ampleur des exigences méthodologiques et techniques de l'évaluation.

Bien que le travail soit à première vue impressionnant, il faut réaliser que les difficultés de répondre aux questions dépendent de la nature des questions. À la question « quelles sont les intentions? », le travail de recherche nécessaire est moins imposant que celui relatif à la question « quels sont les effets? ». Quelques questions font souvent appel aux mêmes méthodes et techniques. Ainsi, la sous-question « quelles sont les cibles de l'intervention? » appelle une instrumentation semblable à celle de la sous-question « quels sont les objectifs de l'intervention? », et la recherche se fera simultanément et fort probablement par les mêmes personnes. Certaines questions, d'ordre normatif, comme « quelle est l'efficacité opérationnelle? » exigent des données colligées auparavant dans le processus d'évaluation, ce qui allège les exigences en temps et en ressources.

**Tableau 2 — Structure typique d'une matrice d'évaluation.**

Démarche d'évaluation	Hypothèse	Variables	Méthode	Indicateurs	Outils	Analyses
<b>Questions d'évaluation</b>						
<i>Étude de la raison d'être</i>						
1. Problématique d'intervention						
2. Justification d'intervention						
<i>Étude de la conception</i>						
3. Intentions (cibles et objectifs)						
4. Nature de l'intervention						
5. Justification des intentions						
6. Justification de la nature de l'intervention						
<i>Étude de la mise en œuvre</i>						
7. Plan de mise en œuvre						
8. Intrants, activité de production, extrants						
9. Conformité au plan de mise en œuvre						
10. Coûts						
11. Efficacité, efficacité et équité opérationnelles						
<i>Étude des effets</i>						
12. Effets						
<i>Étude de la valeur de l'intervention</i>						
13. Efficacité globale						
14. Efficacité globale						
15. Équité globale						
16. Valeur de l'intervention						

Cette série de recherches appliquées est, comme on le sait déjà, caractérisée par une imbrication et une complémentarité élevées des questions d'évaluation. En conséquence, la priorisation des questions ne peut être réalisée efficacement qu'en tenant compte de cette imbrication.

## Des pistes fructueuses pour le choix d'indicateurs

La collecte de données sur le terrain nécessite une instrumentation et des personnes-ressources dont le coût représente souvent la partie la plus importante du budget d'une évaluation. Afin de maintenir les coûts des évaluations à des niveaux réalistes, choisir des indicateurs à partir de bases de données existantes, publiques ou privées, apparaît souvent comme une option attrayante. Les indicateurs et les données de santé publique par exemple, les indicateurs environnementaux, les indicateurs en transport, doivent être considérés lors de l'élaboration de la matrice d'évaluation. Il faut cependant garder à l'esprit qu'il s'agit de données dites « secondaires », qui sont colligées à des fins générales ou à des fins autres que celles de l'évaluation en cours. Ces indicateurs et ces données peuvent servir à établir des tendances générales et des points de comparaison utiles, mais l'évaluation exige la plupart du temps des collectes de données primaires, à partir d'indicateurs spécifiques au projet d'évaluation.

Alors, comment choisir de manière pertinente ces indicateurs et collectes de données tout en réduisant le plus possible le coût des évaluations ?

Il s'agit de concentrer l'effort d'évaluation sur des pistes fructueuses. Et ces pistes existent déjà à l'étape de la construction de la matrice d'évaluation. Elles sont balisées par l'étude de la problématique de l'intervention, des cibles d'intervention, par l'analyse des pour et des contre dans le choix des instruments d'intervention, par l'anticipation des effets grâce aux théories de l'intervention, par les effets hors cibles prévisibles, etc. Pour donner du sens aux indicateurs de l'évaluation et aux données colligées, et pour éviter une dispersion coûteuse, ils doivent être choisis en continuité avec la réflexion initiale sur la raison d'être et la conception de l'intervention. Ces choix doivent être ancrés également dans la connaissance de la mise en œuvre de l'intervention. Bref, le choix des indicateurs utiles à l'évaluation doit se faire sur la base du modèle d'intervention et de l'ensemble des renseignements que son élaboration a nécessité. Afin d'éviter une vision trop centrée sur celle de l'autorité responsable de l'intervention, il faut que cette recherche d'indicateurs demeure ouverte à la découverte terrain. Mais il demeure prioritaire d'ancrer le choix des indicateurs dans la problématique initiale de l'intervention et sa conception. Ainsi, en évitant la dispersion, le projet d'évaluation gagnera en signification et en économie.

Concrètement, les indicateurs sont déduits des variables considérées primordiales lors de l'examen de la problématique, des cibles et des instruments d'intervention. Et aux premiers rangs figurent la théorie du problème et la théorie de l'intervention. En effet, les théories descriptives et normatives proposent naturellement des variables à partir desquelles les indicateurs peuvent être imaginés, et les données, colligées.

### *Le cas de la pollution des eaux*

Prenons le cas d'un problème de pollution des eaux vécu dans un pays donné. La théorie du problème révèle que la pollution des eaux dans ce pays est engendrée par des opérations d'entreprises industrielles, d'entreprises agricoles et également par les rejets de villes. Leur responsabilité juridique des coûts sociaux qu'elles provoquent est variable, selon les lois en cours. La pollution des eaux ne cesse cependant de croître. Il apparaît immédiatement que : 1) la responsabilité des entreprises et des villes



à l'égard de la pollution; 2) la pollution rejetée par ces acteurs, constituent deux variables importantes dans la compréhension du problème et de ses causes. Des données à cet égard devront être colligées pour que les effets d'une intervention visant à contrer la pollution des eaux soient observés et que l'efficacité et l'efficience de l'intervention soient appréciées. Aussi, les variables «responsabilités à l'égard de la pollution» et «pollution rejetée dans les cours d'eau» appellent à choisir des indicateurs adéquats pour concrétiser une collecte de données pertinente qui renseignera sur les liens à priori et à posteriori entre ces deux variables. Ainsi, des indicateurs de niveau de toxicité (métaux lourds, résidus pharmacologiques...), de particules en suspension, de présence d'engrais agricoles ou de contaminants bactériologiques renseigneront sur la production de pollution par les différents acteurs. Le niveau de responsabilité peut se mesurer par les obligations réglementaires ou l'ampleur des redevances à assumer par les différents acteurs. Une intervention nationale ou à l'échelle du bassin versant modifiant le niveau de responsabilité des acteurs pourrait réduire la production de pollution des eaux, et s'enregistrerait grâce aux mêmes indicateurs servant à décrire la situation de pollution à priori, mais réutilisés cette fois à posteriori. L'efficacité, l'efficience et même certaines considérations d'équité de l'intervention pourraient être estimées par ces mêmes indicateurs.

## La méthode comme stratégie de réponse aux questions

Chaque question d'évaluation suscite des besoins variés en observation et en mesure : des variables sont identifiées, des indicateurs sont choisis, quelques fois même plusieurs indicateurs par variable retenue, ce qui nécessite l'emploi de plusieurs outils de collecte de données. L'éparpillement et la dispersion deviennent un danger réel pour une opération coûteuse en temps et en ressources. La nécessité de coordination se fait ainsi jour et c'est là le rôle de la méthode. Ainsi, chaque question d'évaluation doit être associée à une méthode, en d'autres mots à une stratégie de réponse à la question, coordonnant ainsi les divers outils de collecte de données, la collecte de données elle-même et l'analyse de données.

Ainsi, la réponse à la question portant sur les effets d'une intervention doit s'appuyer sur une méthode qui déploiera diverses techniques complémentaires entre elles. Ces techniques, données et analyses apporteront des parties de réponse à la question d'évaluation. La méthode veille à ce que ces parties de réponse forment un tout, satisfaisant les attentes de l'évaluation.

Par exemple, les hypothèses formulées quant aux effets d'une intervention donnent lieu à une liste de variables d'effets et de facteurs externes à l'intervention, pertinents pour l'analyse de données. La méthode prévoit des observations de nature qualitative (par exemple : entrevues, sondages) auprès de la population cible, et des mesures quantitatives (par exemple : taux de glycémie, activité cardiaque...) de l'état de santé d'un échantillon de la population cible. Ces observations et mesures sont réalisées grâce à des outils de collecte de données totalement différents, mais doivent se compléter par l'analyse de données prévue par la méthode. Ce niveau de complémentarité doit le plus possible être planifié lors de la construction de la matrice d'évaluation.

## Le moment propice

La matrice d'évaluation permet la visualisation de chaque étape de la démarche évaluative pour chaque question d'évaluation. Le déploiement de la démarche, de gauche à droite, évoque la logique de construction de la matrice, mais ne révèle pas à quel moment doit se réaliser chacune des opérations de la matrice d'évaluation.

Or, chaque stratégie, chaque technique, a son moment propice. La planification de la recherche de réponse aux questions d'évaluation nécessite de situer dans le temps la conduite des opérations de recherche. Ainsi, une méthode basée sur des groupes contrôles et des groupes exposés à l'intervention peut exiger des mesures à priori et des mesures à postériori. Il peut s'avérer économique de combiner par exemple plusieurs mesures à priori supportant plus d'une question d'évaluation.

La qualité du travail de planification logistique sera déterminante de l'efficacité de l'évaluation et du respect de son budget.

## La communication et la validation

L'élaboration de la matrice d'évaluation met à l'œuvre ou implique de nombreuses parties prenantes. Leur collaboration est incontournable, et les responsables de l'évaluation doivent s'en assurer sans compromettre le recul et l'objectivité que l'on attend de l'évaluation.

L'expérience des évaluatrices et évaluateurs permet certes d'anticiper certains obstacles et éléments facilitateurs à la collaboration, mais l'échange direct avec les parties prenantes ne peut être esquivé. La matrice d'évaluation devient alors un outil de communication fort utile entre les parties par sa transparence et sa précision. Les discussions et les échanges de toute nature (précision et pertinence des questions, hypothèses de réponse, difficultés de réalisation, moments propices pour la collecte de données...) font gagner ainsi en efficacité, permettent aux responsables de l'évaluation de valider les décisions préliminaires et d'effectuer, si nécessaire, des correctifs ou des ajouts.

## Un exemple de matrice d'évaluation

Pour illustrer l'élaboration d'une matrice d'évaluation, nous reprenons ici le cas fictif d'une intervention de lutte contre la pauvreté :

*« Une intervention en matière de lutte contre la pauvreté » (PGLPUR) : cas fictif (Vénétia Sauvain, École nationale d'administration publique, 2020, adapté de l'Initiative nationale pour le développement humain du Royaume du Maroc.)*

### 1. Contexte du cas

*La prise de conscience d'une problématique de pauvreté, s'exprimant souvent sous la forme de précarité, d'exclusion sociale, de chômage, de délinquance et de misère, et touchant particulièrement les femmes et les jeunes, amène l'État à intervenir pour améliorer significativement la situation. Les niveaux ministériel, régional et local sont amenés à agir avec l'intention de diminuer l'ampleur de la pauvreté dans le pays. L'intervention vise à financer des projets d'activité génératrice de revenus (AGR) et de microprojets coopératifs (MPC). Un système de suivi-évaluation de l'intervention est mis en place avec la collaboration des différents paliers d'intervention.*

### 2. Illustration d'une matrice d'évaluation

L'illustration qui suit dans le tableau 3 veut montrer comment on peut appliquer à un cas particulier les principes d'élaboration d'une matrice d'évaluation. Le tableau reprend les grandes questions d'évaluation posées sur le cas du PGLPUR développées au chapitre 10 : *Les questions d'évaluation*. Ces questions d'évaluation et la matrice d'évaluation développées ici peuvent ne pas convenir à des situations nationales ou régionales particulières. Elles veulent uniquement illustrer la manière de construire une matrice d'évaluation. Cet exemple peut s'adapter à différents contextes.

**Tableau 3 — Exemple de matrice d'évaluation : cas fictif.**

DÉMARCHE D'ÉVALUATION DU PGLPUR	Hypothèse	Variables	Méthode	Indicateurs	Outils	Analyses
<b>Questions d'évaluation du PGLPUR</b>						
<i>Étude de la raison d'être du PGLPUR</i>						
1. Problématique d'intervention	La pauvreté bloquerait la contribution des femmes et des jeunes à l'économie.	Pauvreté, chômage, délinquance	Démonstrations causales des liens entre pauvreté, emploi et maux sociaux.	Revenu autonome des femmes, occupations des jeunes, taux de participation à l'emploi, criminalité.	Sondages, entretiens, base de données secondaires.	Corrélations statistiques, analyse de contenu.
2. Justification d'intervention	Le marché du travail à lui seul ne pourrait régler le problème.	Défaillance de marché, intervention de l'État, pauvreté.	Analyse comparative des politiques publiques.	Degré d'intervention de l'État, taux de chômage, % sous seuil de pauvreté.	Bilan des études sur le rôle de l'État dans la lutte à la pauvreté.	Méta-analyse.
<i>Étude de la conception du PGLPUR</i>						
3. Intentions (cibles et objectifs)	Les cibles seraient explicites mais pas les objectifs.	Cibles énoncées sous forme de variables d'emploi, de revenus ou de criminalité; objectifs sous forme de niveau souhaité de ces variables.	Étude des discours et des documents administratifs.	Intentions exprimées dans les débats publics et dans la documentation officielle du PGLPUR.	Collecte des documents officiels et des articles dans les médias, entrevue avec les responsables du PGLPUR.	Analyse de contenu, analyse des associations entre cibles et objectifs.
4. Nature de l'intervention	Les subventions conditionnelles augmenteraient la probabilité de réussite des microprojets stimulant le revenu et réduisant la délinquance.	Taux de subvention par pallier de gouvernement; mobilisation à l'emploi des femmes et jeunes, revenu autonome des femmes, des localités; délinquance.	Revue des théories des subventions, analyse comparative entre nations; analyse des interventions nationales des dernières décennies en matière d'emploi.	% et ampleur des subventions par pallier, conditions des subventions; revenu autonome des femmes, des localités; taux de criminalité.	Recherches théoriques, résultats dans les évaluations d'interventions similaires nationales et internationales.	Analyse documentaire.
5. Justification des intentions	Le choix des cibles du PGLPUR serait appuyé sur la documentation scientifique et la connaissance des situations nationale et régionale.	Qualité de la recension des écrits par les concepteurs, adéquation des cibles avec les recommandations scientifiques et administratives.	Validation par des experts des appuis scientifiques et administratifs au choix des cibles.	Degré de conformité et écarts avec les écrits scientifiques et administratifs; justifications des écarts.	Méthode de Delphes (Delphi).	Analyse des convergences et des consensus.
6. Justification de la nature de l'intervention	Le choix de l'instrument d'intervention du PGLPUR serait appuyé sur la documentation scientifique et la connaissance des situations nationale et régionale.	Qualité de la recension des écrits, adéquation de la subvention conditionnelle avec les recommandations scientifiques et administratives.	Validation par des experts des appuis scientifiques et administratifs au choix de l'instrument d'intervention.	Degré de conformité et écarts avec les écrits scientifiques et administratifs; justifications des écarts.	Méthode de Delphes (Delphi).	Analyse des convergences et des consensus.

## L'évaluation en contexte de développement

DÉMARCHE D'ÉVALUATION DU PGLPUR	Hypothèse	Variables	Méthode	Indicateurs	Outils	Analyses
<i>Étude de la mise en œuvre du PGLPUR</i>						
7. Plan de mise en oeuvre	Le plan serait explicite, détaillé et bien diffusé.	Cibles opérationnelles, objectifs opérationnels.	Recherche de la documentation sur le plan de mise en oeuvre.	Énoncés des attentes en termes d'intrants, de l'activité de production et des extrants.	Collecte de la documentation administrative, entrevue avec décideurs clés.	Analyse documentaire.
8. Intrants, activité de production, extrants	Les intrants, l'activité de production et les extrants sont bien documentés dans les rapports de gestion et dans le suivi-évaluation.	Montant et taux des subventions, mécanisme d'attribution des subventions, montants versés.	Centrée sur la description.	Conditions d'attribution, montants en subventions désagrégés par projet et par catégories.	Techniques comptables, classification des conditions selon la théorie économique du secteur public.	Analyses opérationnelles.
9. Conformité au plan de mise en oeuvre	La mise en oeuvre serait conforme au plan de mise en oeuvre au niveau national mais de nombreux écarts (possiblement justifiés) seraient constatés au niveau des régions.	Montant et taux des subventions, mécanisme d'attribution des subventions, montants versés.	Comparaisons entre le plan de mise en oeuvre et les observations et mesures sur la mise en oeuvre.	Cibles et objectifs opérationnels, ressources, opérations, produits.	Utilisation des données de suivi-évaluation quantitatives et qualitatives.	Analyse de type audit.
10. Coûts	Les coûts seraient observables dans les données et rapports comptables de suivi-évaluation du PGLPUR.	Coûts totaux de production et coûts unitaires de production.	Méthodes comptables.	Selon les données recueillies pour le suivi-évaluation, ventilation pour chacun des extrants.	Techniques comptables.	Analyses comptables.
11. Efficacité, efficacité et équité opérationnelles	La mise en oeuvre du PGLPUR serait efficace, efficiente et équitable tant au niveau national qu'au niveau régional.	Efficacité opérationnelle, efficacité opérationnelle, équité opérationnelle.	Comparaison des objectifs aux extrants, comparaison des coûts unitaires de production à d'autres programmes et entre régions; comparaison des montants octroyés par région.	Degré d'atteinte des objectifs opérationnels, coûts unitaires de production, distribution régionale des fonds	Techniques quantitatives et qualitatives variées.	Analyses comparatives.
<i>Étude des effets du PGLPUR</i>						
12. Effets	Les effets seraient produits selon les prévisions (effets prévus). Ils seraient répartis suivant la chaîne de cibles. Peu d'effets hors cible auraient été produits.	Cibles directes, intermédiaires, ultimes, situations hors cible.	Devis quasi-expérimental sans groupe contrôle.	Revenu autonome des femmes, occupations des jeunes, taux de participation à l'emploi, taux de criminalité.	Mesures à priori et à postériori des indicateurs de l'état des cibles et des situations hors-cible.	Calcul de l'effet net du PGLPUR sur chaque indicateur.

DÉMARCHE D'ÉVALUATION DU PGLPUR	Hypothèse	Variables	Méthode	Indicateurs	Outils	Analyses
<i>Étude de la valeur du PGLPUR</i>						
13. Efficacité globale	N/A. L'évaluation est en mesure de rendre compte des effets du PGLPUR mais l'absence d'objectifs rend impossible la lecture de l'efficacité.	S.O. à cause de l'absence d'objectif.	S.O. à cause de l'absence d'objectif.	S.O. à cause de l'absence d'objectif.	S.O. à cause de l'absence d'objectif.	S.O. à cause de l'absence d'objectif.
14. Efficience globale	Le PGLPUR présenterait un rendement des ressources plutôt modéré quant à la réduction de la délinquance mais relève substantiellement le revenu autonome des femmes en situation de pauvreté.	Revenu des femmes et délinquance des jeunes en milieu de pauvreté, coût total de production du PGLPUR, autres avantages et coûts sociaux.	Analyse de rendement des ressources.	Revenu autonome des femmes, occupation des jeunes, taux de participation à l'emploi, taux de criminalité.	Technique coût-avantage et méthodes pondérées.	Analyses quantitatives et qualitatives selon la technique retenue.
15. Équité globale	Redressement significatif de la situation économique des femmes pauvres par rapport à d'autres classes sociales.	Revenu des femmes et délinquance des jeunes selon les classes sociales.	Évolution des écarts selon les classes sociales.	Revenu autonome des femmes, occupation des jeunes, taux de participation à l'emploi, taux de criminalité.	Techniques de comparaison quantitatives et qualitatives.	Analyse des écarts et de la signification statistique des écarts lorsque possible.
16. Valeur de l'intervention	Malgré l'absence de lecture de l'efficacité, le PGLPUR se montrerait modérément efficient comparativement à d'autres interventions et corrigerait sensiblement des situations socialement inéquitables.	Efficience globale et équité globale.	Jugement quant à l'équilibre entre l'efficience globale modérée mais une équité globale significative.	S.O.	Méthode pondérée.	Explicitation de la pondération des critères d'efficience globale et d'équité globale.

Le tableau 3 montre comment l'évaluation d'une intervention peut adopter une démarche de type « recherche appliquée ». L'élaboration de la matrice d'évaluation exige de nombreuses décisions typiques de la recherche appliquée. La suite des étapes d'évaluation peut alors se réaliser grâce à une logistique précise et coordonnée. L'estimation des coûts de l'évaluation peut se faire sur une base explicite, et la décision d'enclencher les opérations subséquentes pourra se prendre en connaissance de cause.

Cet exemple de matrice d'évaluation fait également ressortir dans toute sa dimension les exigences méthodologiques et techniques de l'évaluation. Cette vision d'ensemble des travaux évaluatifs doit être gardée à l'esprit des responsables de l'évaluation lors des décisions prises sur les questions d'évaluation. D'une part, chaque question ajoutée implique une démarche évaluative et des coûts logistiques supplémentaires. D'autre part, tout retranchement des questions présentées au tableau 3 a des implications sur la capacité de répondre à des questions subséquentes.

## Conclusion

Ce chapitre sur la matrice d'évaluation a permis de montrer que pour obtenir des réponses aux questions d'évaluation, il y a lieu d'adopter une démarche de type recherche appliquée. Il explique que la matrice d'évaluation se présente sous une forme « matricielle » parce qu'une évaluation implique plusieurs questions d'évaluation et que chaque question d'évaluation doit disposer de sa propre démarche de réponse. Concrètement, à chacune des questions, disposées en colonne, correspond une ligne particulière qui précise les éléments de la démarche. Ainsi les hypothèses, les variables, la méthode, les indicateurs, les outils de collecte de données et les analyses de données sont présentés pour chaque question d'évaluation. Chaque démarche de réponse possédant son « moment propice », la matrice d'évaluation se déploie de manière ordonnée dans le temps, avec un calendrier établi. L'élaboration de la matrice d'évaluation est l'occasion de concrétiser les priorités dans les réponses recherchées et de réaliser l'arbitrage entre la qualité des réponses et les coûts impliqués pour l'évaluation. La communication et l'échange entre les responsables de la matrice d'évaluation et les parties prenantes sont essentiels à la réalisation efficace et au succès de l'évaluation. La validation de la matrice d'évaluation par des instances compétentes lui confère une crédibilité indispensable.

## Messages clés

1. Pour obtenir une ou des réponses à une question d'évaluation, il faut une démarche qui lui est propre.
2. Pour qu'une réponse à une question d'évaluation soit de qualité, il faut adopter une démarche de type recherche appliquée.
3. La matrice d'évaluation présente une forme « matricielle » parce qu'une évaluation implique plusieurs questions d'évaluation et que chaque question dispose de sa propre démarche. Concrètement, à chacune des questions, disposées en colonne, correspond une ligne qui précise les éléments de sa démarche.
4. Chaque démarche de réponse possédant son « moment propice », la matrice d'évaluation se déploie de manière ordonnée dans le temps, avec un calendrier établi au préalable.
5. Le développement de la matrice d'évaluation est l'occasion de concrétiser les priorités dans les réponses recherchées et de réaliser l'arbitrage entre la qualité des réponses et les coûts impliqués pour l'évaluation.
6. La communication et l'échange entre les responsables de la matrice d'évaluation et les parties prenantes sont essentiels à la réalisation efficace et au succès de l'évaluation.
7. La validation de la matrice d'évaluation par des instances compétentes lui confère une crédibilité indispensable.

## Exercice de réflexion

- L'organisation pour laquelle vous travaillez soumet quelques questions supplémentaires à propos du PGLPUR, questions suggérées lors d'une consultation des parties prenantes. Devez-vous ajouter ces questions à la matrice d'évaluation? Si non, pourquoi? Si oui, comment ferez-vous le tri et à quelle ligne de la matrice devez-vous les placer?
- Supposons qu'une de ces nouvelles questions porte sur la durabilité de l'intervention. Quelle sera votre approche dans la gestion de la matrice d'évaluation?
- Une partie prenante influente anticipe une conséquence dommageable du programme, conséquence non incluse dans la matrice d'évaluation. Comment procéderez-vous?
- Votre organisation souhaite qu'il y ait, malgré l'absence d'objectifs, une analyse de l'efficacité du PGLPUR. Comment réagirez-vous?
- Avant même de débiter la recherche terrain, on vous signifie que le budget de l'évaluation est réduit du tiers. Comment procéderez-vous pour respecter le nouveau budget? Quelles seront les priorités? Comment tiendrez-vous compte de l'imbrication des questions les unes dans les autres?

## Pour aller plus loin

### 1. Cas du Programme d'aide juridique du Gouvernement du Canada.

Ce lien conduit à la table des matières du rapport entier d'évaluation. L'annexe B, accessible, présente la matrice d'évaluation utilisée pour l'évaluation. Les questions d'évaluation, listées typiquement en colonnes, prennent une forme spécifique associée à la nature du programme lui-même et à la terminologie particulière à la *Politique sur l'évaluation* du gouvernement canadien. La démarche d'évaluation, présentée en ligne, contient trois éléments : questions, indicateurs possibles et sources de données. Bien que le spectre de la matrice soit plus limité que celui plus général proposé dans ce chapitre, ce qui rend compte, au passage, du contexte de réalisation de l'évaluation, la démarche d'évaluation est clairement de type recherche appliquée. Ce cas montre avec quels soins et détails la matrice d'évaluation a été conçue.

<https://www.justice.gc.ca/fra/pr-rp/pm-cp/eval/rap-rep/2016/paj-lap/tdm-toc.html>

### 2. Matériel de formation du Fonds de l'OIM pour le développement (ONU Migration, évaluation et suivi).

Ce lien conduit à du matériel de formation apparenté au contenu de ce chapitre. Vous y trouverez une « matrice de résultats ». Les colonnes réfèrent implicitement à des questions relatives aux résultats tels que les objectifs, les effets directs, les produits et les activités, alors que les lignes présentent le contenu de la démarche évaluative proposée dans un ordre cependant différent de celui de ce chapitre. Cette variété dans la manière de présenter la matrice d'évaluation peut déstabiliser en début d'apprentissage, mais avec la pratique et l'approfondissement des connaissances en évaluation, les praticiens

de l'évaluation s'en formalisent moins. Nous référons les lecteurs et lectrices à une lecture attentive des chapitres précédents de ce manuel ainsi qu'au glossaire de chapitre 2 : *La terminologie de l'évaluation* pour assurer la cohérence dans les concepts mêmes de l'évaluation et la manière de l'appliquer

<https://www.iom.int/sites/g/files/tmzbdl486/files/about-iom/evaluation/evaluation-criteria-toolkit-fr.pdf>

3. Cas du Guide de la planification, du suivi et de l'évaluation axés sur les résultats du développement du Programme des Nations unies pour le développement (PNUD).

La section 3.2 du Guide du PNUD intitulée *Le Cadre de suivi et d'évaluation* présente le tableau 15 : *Exemple de Matrice de planification du suivi : capacité renforcée de l'autorité administrant les élections*. La démarche de construction de la matrice est présentée dans la première ligne et contient essentiellement les mêmes éléments que ceux présentés dans ce chapitre. Au demeurant, ce tableau fait la démonstration que cette démarche peut être précisée de différentes façons selon les institutions pour lesquelles l'évaluation est requise. Ici encore, les questions présentées dans la première colonne sont centrées exclusivement sur les résultats. Cela montre la priorité donnée par l'institution aux réponses aux questions portant sur les résultats. Ce cadre peut cependant être mis en question quant au réalisme d'une évaluation qui n'aurait pas au préalable tenté de répondre aux questions relatives à la problématique d'intervention, à la conception de l'intervention et à sa mise en œuvre. Une hypothèse de travail à considérer est que d'autres composantes de l'institution ont déjà tenté de répondre aux questions préalables aux résultats.

<http://web.undp.org/evaluation/handbook/french/ch3-2.html>



# Choix des outils pour les stratégies de collecte et d'analyse des données en évaluation



Vénétia Sauvain  
Ousséni Kinda  
Maxime Desfossés

## Introduction

Comme cela a été abordé dans le chapitre précédent, face à la diversité des données ou des informations possibles pour répondre aux questions d'évaluation, l'évaluateur doit mettre en place une stratégie de collecte et d'analyse des données à la fois pragmatique et qui se rapproche le plus possible des critères de rigueur scientifique, qu'on appellera « normes », pour éviter de les confondre avec les critères d'évaluation. En effet, puisque l'évaluation emprunte aux procédures issues des sciences sociales dans le cadre d'une démarche systématique de collecte et d'analyses des données (Rossi *et al.*, cité par Delahais *et al.*, 2021), ces règles de jugement servent de fil conducteur pour la mise en œuvre de combinaisons, de méthodes et d'outils quantitatifs et qualitatifs. Ce chapitre présente d'abord de manière synthétique ce qui caractérise la nature des stratégies de collecte et d'analyse de données (quantitative, qualitative, mixte) et les principales normes qui encadrent le choix et la mise en œuvre d'outils de collecte. Ce choix découle naturellement des approches évaluatives adoptées<sup>1</sup>. Il est également déterminé par les apports et les limites méthodologiques inhérentes à ces outils, par l'envergure de l'évaluation (en lien avec les délais et le budget accordé) ainsi que par la disponibilité et l'accessibilité des sources. Dépendamment des cultures organisationnelles et des finalités de l'évaluation, certains outils seront utilisés plus souvent que d'autres. Ce chapitre n'a pas la prétention d'être exhaustif sur les outils possibles et imaginables dans le domaine. Il n'entre pas non plus dans les détails des applications et des limites méthodologiques de chacun. Concernant les outils classiques de collecte, la vaste littérature en français dans les sciences économiques et sociales est suffisante pour ceux et celles qui aimeraient approfondir leurs connaissances en la matière. Toutefois, ce chapitre accorde une attention particulière aux outils participatifs de collecte de données, avec des illustrations de leur application, pour contribuer à la mise en œuvre du principe « aucun laissé pour compte » : comment prendre en compte les perspectives des plus vulnérables à travers le choix des outils ?

1. Cf. le chapitre 7 : *Le choix des approches évaluatives*.

## Stratégie de collecte et d'analyse des données

Comme cela a été mentionné dans le chapitre précédent, « la matrice d'évaluation permet de rassembler et de coordonner la logistique évaluative en un seul tableau pratique » (Marceau, 2022). La stratégie de collecte et d'analyse des données permet d'optimiser les choix de la matrice<sup>2</sup>. Quelle que soit la nature de la stratégie retenue, **la rigueur** et **l'authenticité** sont deux qualités inhérentes à la démarche (Fortin et Gagnon, 2016), et cette dernière se doit d'être transparente. Les critères pour juger de la rigueur et de l'authenticité sont toutefois différents selon que la stratégie est de nature qualitative ou quantitative, car chacune repose sur des paradigmes différents. Avant d'expliquer les caractéristiques, avantages et limites pertinents à chacune des stratégies, il est important de rappeler quelques éléments propres à la nature des données utilisées dans les évaluations.

### La nature des données utilisées dans les évaluations

Les données correspondent aux informations et aux observations qui sont collectées, traitées, analysées et interprétées pour répondre aux questions d'évaluation. Il existe des données numériques (dites quantitatives) et des données textuelles ou narratives (dites qualitatives). Certaines approches d'évaluation utilisent également des données visuelles (images, vidéos), géospatiales ou sonores. Les données d'évaluations sont d'abord de sources **secondaires**, comme il a été mentionné dans le chapitre précédent. Selon notre expérience, la majorité des données secondaires utilisées dans les évaluations sont : des données administratives des organisations publiques, la plupart du temps provenant d'un ministère, et correspondant à des indicateurs d'extrants (par exemple, le nombre d'écoles, classées par régions administratives, le nombre d'élèves dans l'école selon des critères d'âge et de sexe, le taux de scolarisation par région, etc.) ou de résultats d'effets directs (le taux de réussite à l'école, le taux de diplomation); des données de suivi des interventions de développement (selon des indicateurs de cadre logique ou de cadre de mesure et de rendement); ou encore des données statistiques issues des grands indicateurs macroéconomiques ou nationaux (par exemple, le taux de chômage, la couverture médicale, le taux de scolarité, etc.) qui permettent d'informer les indicateurs des effets ultimes. La plupart des données secondaires mentionnées jusqu'à présent sont plutôt de nature quantitative, il s'agira alors de faire des analyses quantitatives pour les fins de l'évaluation. Généralement, ces données peuvent être dispersées ou sous une forme « brute » (c'est-à-dire non analysée, par exemple dans un fichier Excel) et le travail de l'équipe d'évaluation pourra être de les traiter, de les colliger, etc. aux fins de l'évaluation. Beaucoup de ces données pourront être traitées au préalable et seront utiles aux décisions de collectes subséquentes ou pourront aider au choix d'échantillonnage. Dans les évaluations axées sur la pertinence ou la cohérence de l'intervention, l'enjeu sera plutôt d'établir les intentions des acteurs, de caractériser le problème ou la situation à l'origine de l'intervention. Certains éléments en lien avec ce problème seront déjà disponibles sous forme de données secondaires, par exemple, à travers des revues de littérature ou la publication de rapports thématiques sur le problème. Il s'agira donc plutôt d'analyser le contenu de ces documents. Cela dit, une évaluation ne pourra pas se contenter

---

2. Comme cela a été démontré dans l'exemple de matrice du chapitre précédent, un même outil auprès d'une même source peut renseigner sur plusieurs indicateurs/variables à la fois. Il peut être utilisé plusieurs fois auprès de différentes populations. Il peut être mobilisé par une méthode A et une méthode B, etc.

des sources secondaires, car elle nécessite de faire le lien entre l'environnement de l'intervention et celle-ci, quels que soient la nature et l'objet des questions d'évaluation, ce qui nécessite généralement de collecter de l'information supplémentaire. C'est dans ce contexte qu'on parle alors d'outils<sup>3</sup>.

### Ce qui caractérise les méthodes quantitatives et leurs normes de qualité

Une stratégie quantitative est une stratégie de collecte et d'analyse de données qui regroupe des méthodes et des outils quantitatifs (p. ex. des sondages (questions fermées), des recensements ou des enquêtes de ménages, des analyses de tendances, etc.). Elle permet de collecter un nombre important d'observations pour décrire de manière chiffrée une situation précise ou d'établir une relation de cause à effet entre des variables choisies. Elle met « l'accent sur la description, l'explication, la prédiction et le contrôle et repose sur la mesure de phénomènes<sup>4</sup> et l'analyse de données numériques » (Fortin et Gagnon, 2016, p. 29). Elle s'inscrit dans les paradigmes positivistes et postpositivistes. Une stratégie quantitative met en œuvre des techniques d'échantillonnage et des outils de collecte et d'analyse de données qui permettent de faire des liens entre les variables (p. ex. estimer et déterminer si l'adoption d'un comportement particulier dépend d'une caractéristique des individus, ou si elle est liée à un environnement particulier, ou à l'adoption d'autres comportements en parallèle, etc.), ainsi que la généralisation des constats.

Plusieurs normes de qualité guident le raisonnement hypothético-déductif propre aux méthodes quantitatives, notamment les éléments de fiabilité (qualité de la mesure « en soi ») et de validité (la représentation adéquate du concept) qui déterminent la qualité de la mesure, c'est-à-dire des indicateurs associés aux variables (Durand, 2021a). Au niveau de la « structure de la preuve » (Gauthier, 2021, p. 130), c'est-à-dire le raisonnement qui permet d'établir des constats valides, par exemple la logique de démonstration qui permet d'établir un jugement sur les effets d'une intervention, on peut retenir la validité interne et la validité externe. La validité interne correspond à « la mesure dans laquelle une évaluation permet d'établir efficacement le lien de cause à effet d'une intervention sur les résultats d'intérêt » (Unité de l'impact et de l'innovation, 2019). La validité externe correspond au fait de pouvoir généraliser les constats à d'autres populations ou contextes que ceux étudiés. Elle renvoie à la notion de représentativité des échantillons (Fortin et Gagnon, 2016). Ces deux normes font des méthodes quantitatives les méthodes privilégiées des évaluations des effets qui sont présentées dans le prochain chapitre. Le contrôle de la validité interne permet en effet d'établir un lien clair entre l'intervention et le changement observé. Toutefois, l'information est alors limitée à quelques variables (celles identifiées à l'avance, selon les hypothèses de la personne évaluatrice) et on ne possède généralement que peu d'informations pour comprendre pourquoi et comment le changement a été obtenu. Lorsqu'il s'agit d'estimer le changement auprès d'une grande population exposée (plusieurs

3. Plusieurs termes sont utilisés pour qualifier les choix méthodologiques qui sont effectués en évaluation. Parfois, « méthodes » et « outils » sont utilisés comme synonyme; parfois c'est « approches » et « méthodes ». Parfois, « *design* » en anglais est utilisé comme équivalent de devis (plutôt dans le domaine de la santé publique), parfois comme équivalent à méthode, stratégie, procédé, protocole... Il est de notre avis que peu importe la culture académique et organisationnelle de chacun, l'important est de bien savoir communiquer ce que l'on signifie aux personnes impliquées dans l'évaluation. De notre côté, nous associons le mot « outil » aux instruments et techniques qui permettent concrètement de collecter des observations et de faire des mesures. Les méthodes mobilisent donc plusieurs outils selon un protocole établi qui fait office de méthodologie. La combinaison ou pas de ces méthodes reposent sur la notion de paradigme ou courant de pensée, c'est-à-dire comment on pense que l'on peut représenter le monde/le réel. Cf. Dagenais et Ridde, (2012), aussi Ridde *et al.* (2021), à propos de la notion de paradigme en évaluation.

4. Le phénomène renvoie à ce qui est observé.

milliers de personnes), elles permettent d'estimer la magnitude et la distribution des impacts (ce qui n'est pas toujours nécessaire pour les évaluations portant sur des petites interventions où la population cible est de moins d'une centaine). Les menaces ou les obstacles à la validité interne et externe sont présentés dans l'encadré.

**Encadré 1 – Menaces à la validité interne et externe (adapté de Gauthier, 2021, p. 144-147 et Anstett, 2016)**

Menace à la validité interne	Menace à la validité externe
<p><b>Facteurs historiques</b> : Des événements externes ont lieu en même temps que l'intervention, et ont potentiellement influencé les changements observés sur la variable.</p> <p><b>Maturation</b> : Les groupes étudiés changent par eux-mêmes par l'effet du temps sans lien avec l'intervention.</p> <p><b>Fluctuation de l'instrument de mesure</b> : La mesure ne se fait pas de manière constante.</p> <p><b>Apprentissage au test</b> : La répétition de la mesure fausse la mesure.</p>	<p><b>Biais de sélection, non-représentativité</b> : L'échantillon ne reflète pas adéquatement la population.</p> <p><b>Contagion</b> : L'intervention sur un groupe a un effet sur d'autres groupes, car les groupes ne sont pas étanches.</p> <p><b>Mortalité expérimentale</b> : Perte des sujets durant l'intervention.</p> <p><b>Réactivité aux conditions expérimentales</b> : Il arrive que les sujets, se sachant évalués, agissent ou répondent en fonction de ce qui est attendu d'eux (p. ex. par désir de plaire), ou en réaction à la mesure initiale.</p>

La démarche d'une stratégie quantitative est **structurée**, c'est-à-dire que toutes les données sont collectées exactement de la même manière. Les questions des questionnaires sont élaborées à l'avance et ne peuvent être changées. Les questions posées sont fermées, c'est-à-dire qu'un choix de réponses est défini et proposé aux répondants. Les techniques d'administration du questionnaire sont également convenues à l'avance et encadrées. La technique d'échantillonnage est déterminée à l'avance et vise la représentativité de la population.

### Ce qui caractérise les méthodes qualitatives et leurs normes de qualité

Une stratégie qualitative est une stratégie de collecte et d'analyse de données qui regroupe des méthodes et des outils qualitatifs (des entrevues (questions ouvertes), des groupes de discussion, des observations, la cartographie des communautés, etc.). Elle permet de collecter de l'information **de manière approfondie**. Les sources d'information sont multiples, mais le nombre de personnes est restreint. Elle vise une connaissance en profondeur d'un phénomène, elle se concentre sur le vécu des individus et le sens qu'ils donnent à leurs expériences. Les méthodes qualitatives sont privilégiées en évaluation pour bien comprendre les intentions des acteurs, la dynamique et les processus des acteurs et de l'intervention, ainsi que pour donner du sens au « succès » ou à « l'échec » d'une intervention. Une stratégie qualitative privilégie **une démarche semi-structurée**, c'est-à-dire que la stratégie s'adapte au fur et à mesure de l'émergence des données, de leurs analyses et de leur interprétation. Les techniques et les outils mis en œuvre privilégient des questions ouvertes (sans choix de réponses prédéterminés) pour obtenir de l'information auprès des répondants en étant le plus prêt de leur ressenti, de leurs opinions

et de leurs expériences. La taille de l'échantillon est raisonnée, elle privilégie les personnes ou les sources à haute teneur en information, dépend des données recueillies au fur et à mesure et vise à atteindre « la saturation des données », c'est-à-dire le moment dans la collecte où on peut conclure qu'une nouvelle information n'ajoutera rien de plus à la compréhension de ce qui est étudié (Fortin et Gagnon, 2016).

Parmi les avantages d'une stratégie qualitative en évaluation, on peut citer : la flexibilité de la démarche, l'obtention de constats plus nuancés et approfondis, la possibilité de mettre en valeur des éléments plus proches du terrain au moment de la présentation des résultats (des « verbatim » significatifs, des récits imagés, des descriptions fécondes) ce qui les rend plus tangibles et accessibles pour le lecteur de rapport. L'un des avantages essentiels de compter sur des méthodes qualitatives dans une stratégie de collecte et d'analyse des données en évaluation est de pouvoir prendre le pouls du ressenti des parties prenantes et des populations cibles ou exposées, allant jusqu'à les rendre acteurs des constats d'évaluation lorsqu'elles sont intégrées dans l'ensemble de la démarche d'évaluation selon les approches participatives. Les méthodes qualitatives, et participatives, qui prennent leur racine dans l'enquête de terrain (et qui seront abordées dans un autre chapitre) sont essentielles pour obtenir de l'information auprès des populations « laissées pour compte », défavorisées. En effet, les éléments de contrôle des enquêtes auprès des populations (c'est-à-dire les sondages en présentiel), indispensables à l'obtention de données rigoureuses, deviennent des obstacles lorsqu'il s'agit d'atteindre des populations défavorisées. Prenons un exemple très simple, une bonne question de sondage fermée, c'est-à-dire qu'elle propose un choix de réponses déterminé. Bien qu'elle doive être claire et le plus proche du vocabulaire du répondant, elle reste directe et ne laisse pas beaucoup de place à une réponse nuancée ou contextualisée; ce qui est d'autant plus maladroit dans les cultures marquées par l'oralité, habituées aux périphrases, etc. De plus, l'administration à l'oral de tout un sondage est marquée par les inattentions, le jeu, l'incompréhension, l'ennui. Et lorsqu'on est face à des personnes montées contre un projet, un programme, un acteur, etc., elle ne permet pas l'épanchement que permet une entrevue, même semi-dirigée.

Toutefois, l'ensemble des techniques essentielles à maîtriser pour mettre en œuvre de manière rigoureuse les méthodes qualitatives ne sont pas nécessairement courantes et s'acquièrent avec la pratique (contrairement aux méthodes de formation rapides et standardisées qui accompagnent les personnes dépêchées pour faire passer des sondages). Ces techniques permettent de limiter les biais des individus qui collectent l'information (identifié comme les « biais du chercheur ») et ceux des répondants, c'est-à-dire la rétention d'information, les informations délibérément erronées, la mémoire sélective, l'interprétation des questions ou d'événements. Ces techniques sont essentielles pour respecter les normes de qualité qui guident la stratégie de collecte et d'analyse qualitative, et qui ont été proposées par Lincoln et Guba (1985), comme pendant à la stratégie quantitative : la crédibilité, la transférabilité, la fiabilité et la confirmabilité. La **crédibilité**, miroir de la validité interne, est le critère qui détermine « dans quelle mesure la description du phénomène vécu par les participants reflète la réalité interprétée » (Fortin et Gagnon, 2016, p. 377). La **transférabilité**, miroir de la validité externe, vise à utiliser les constats propres à un contexte pour tirer des leçons valables pour d'autres contextes. La **fiabilité**, équivalente à la fiabilité de l'indicateur, correspond à l'idée que la répétition de l'étude avec les mêmes personnes dans des circonstances similaires donnerait les mêmes constats. La **confirmabilité** permet de juger de l'intégrité de l'étude qualitative en lien avec la diversité des sources, la neutralité des données collectées et leur interprétation en limitant les biais du chercheur (Fortin et Gagnon, 2016).

## L'échantillonnage

La mise en œuvre d'une stratégie de collecte et d'analyse des données nécessite de procéder au choix des individus auprès de qui on souhaite collecter de l'information. En évaluation, on peut faire le choix d'étudier une population (l'ensemble social que l'on souhaite étudier, qui peut être « un ensemble de personnes, d'organisations ou d'objets de quelque nature que ce soit » ou « ensemble d'éléments constituant un tout » (Campenhoudt *et al.*, 2017, p. 209-210) en entier, car elle est relativement petite et qu'il s'agit d'un nombre de sujets auxquels on a facilement accès (par exemple l'ensemble des chefs d'un corps de police qui pourraient recevoir un sondage électronique transmis par leur chef de corps, lui-même notifié par le ministère de l'Intérieur). On peut aussi utiliser des données de sources secondaires portant sur des populations, comme les données de recensement par exemple. Cela dit, il n'est pas toujours nécessaire d'étudier une population dans son entier. Il est possible de sélectionner les « unités » qui constituent la population selon des techniques d'échantillonnage qui permettront d'avoir une bonne représentativité de la population. Typiquement, un millier d'unités interrogées suffit pour constituer un sondage d'opinion par rapport à une population de plusieurs millions (au-delà d'une population de 20 000 unités, il n'est plus nécessaire d'augmenter la taille de l'échantillon). Parmi chaque population, il s'agira alors de sélectionner des unités selon les techniques d'échantillonnage appropriées pour assurer la représentativité de l'échantillon, c'est-à-dire « que les caractéristiques mêmes de la population soient présentes dans l'échantillon ou puissent y être retrouvées moyennant certaines modifications » (Beaud, 2021, p. 191). L'échantillonnage aléatoire consiste à choisir les unités parmi la population, de telle sorte que chacune ait la même chance d'être sélectionnée, ce qui permet de garantir la représentativité de l'échantillon par rapport à la population plus large. Il est utilisé lorsque la population est importante et que beaucoup de données collectées sont envisagées pour chaque unité, et lorsqu'on souhaite que les informations obtenues auprès de ces unités soient globalement les mêmes que si on avait interrogé l'ensemble de la population. Plusieurs techniques d'échantillonnage aléatoire existent : aléatoire simple, échantillon systématique (individus sélectionnés à intervalle fixe à partir d'une liste) ou aréolaire (sélection à partir d'une carte géographique, d'une photo ou d'un plan), en grappes (tirage au sort de groupes d'individus), ou stratifié (division de la population à étudier en sous-populations appelée « strates », tirage aléatoire dans chacune de ces strates) (Beaud, 2021). Chacune possède ses avantages et ses inconvénients pour assurer la représentativité de l'échantillon. L'enjeu des différentes techniques d'échantillonnage correspond surtout à la possibilité de les mettre en œuvre en fonction de l'information que l'on détient pour rejoindre la population, des informations caractéristiques de la population dont on dispose pour établir des quotas ou encore de la mise en œuvre de la collecte<sup>5</sup> (Campenhoudt *et al.*, 2017). En évaluation, le client principal devra être en mesure de fournir l'information correspondant à l'accessibilité de la population (qu'elle soit entière ou sélectionnée) pour pouvoir la rejoindre. Tout dépend alors des réglementations en vigueur dans les pays où ont lieu ces enquêtes sur la confidentialité des données (p. ex. pour un sondage électronique, les milieux de travail organisés devraient disposer de registres de coordonnées professionnelles, mais il est bien plus difficile de rejoindre une population qui ne dispose pas de courriels professionnels). Le questionnement (dans notre cas, les questions d'évaluation), les caractéristiques de la population étudiée et les différentes contraintes logistiques et financières orientent le choix des techniques d'échantillonnage (Beaud, 2021). Selon les contextes, des techniques non probabilistes peuvent être plus adaptées.

5. Cf. le chapitre « L'entrée sur le terrain en évaluation : enjeux et pratiques des collectes de données » pour avoir un aperçu des problématiques de ce dernier aspect.



On peut également opter pour l'étude d'un échantillon non probabiliste, c'est-à-dire dont les composantes sont « non strictement représentatives de la population, mais *caractéristiques* de la population » selon un choix raisonné (Campenhoudt *et al.*, 2017, p. 212). C'est le cas de la sélection de cas. Une étude de cas peut être définie comme « une approche de recherche qui consiste à enquêter sur un phénomène, un événement, une organisation ou un groupe d'individus bien délimité, afin d'en tirer une description précise et une interprétation qui dépasse ses bornes » (Roy, 2021, p. 150). Dans le contexte de l'évaluation d'une intervention, le cas correspond généralement à l'organisation ou à la communauté cible de l'intervention ou encore à quelques bénéficiaires choisis pour raconter les changements qu'ils observent. Par exemple, si l'intervention consiste à destiner à une trentaine de communautés des fonds pour leur développement économique étendu sur tout le territoire national, il est fort probable qu'il soit impossible de visiter chacune de ces communautés. Une sélection de communautés devra s'opérer selon certains critères afin de tirer des enseignements en fonction du contexte de chaque communauté. Les communautés sélectionnées deviendront ainsi des cas. L'étude de cas permet la collecte du plus grand nombre d'informations possible sur chaque cas; alors que les études d'échantillon ont nécessairement besoin d'un grand nombre de sujets, mais ne peuvent recueillir qu'un nombre limité d'informations. Dépendamment des contextes, les études de cas sont utilisées comme des approches d'évaluation à part entière : l'évaluation s'opérationnalise uniquement à travers l'étude d'un ou de plusieurs cas. La matrice d'évaluation s'appliquera alors à chacun de ces cas. Dans d'autres contextes, il pourra être décidé que seulement certaines questions d'évaluation feront l'objet d'une étude de cas. Le choix des cas est principalement déterminé par la capacité de l'équipe d'évaluation à obtenir le plus d'informations possible (généralement dans un court laps de temps) et à pouvoir tirer des comparaisons intéressantes entre les cas. Il est important de clarifier les attentes du client principal de l'évaluation sur la sélection des cas. En effet, le choix d'un seul et unique cas extrême peut être riche en « leçons », en lien avec des problématiques particulières, mais celles-ci peuvent elles-mêmes limiter les enseignements plus généraux. Le choix de cas multiples peut permettre de comparer l'efficacité et l'efficience d'une intervention. On peut alors préférer une sélection de cas homogènes du point de vue des contextes pour limiter les facteurs explicatifs qui sont liés au contexte, et dégager des tendances générales de nature descriptive. Au contraire, on peut viser l'hétérogénéité des cas pour mieux comprendre comment les contextes influencent la manière dont l'intervention s'opérationnalise dans la réalité et comment le changement s'opère (ou comment sont obtenus les effets). Selon notre point de vue, les études de cas prennent naissance d'une volonté d'explication et de compréhension approfondie de ce qui fonctionne et de ce qui ne fonctionne pas. D'autres outils permettront de dégager des tendances générales (analyses de base de données et sondages). Ainsi, si on en revient à notre choix de communauté, on peut partir de l'analyse d'une base de données qui permet de constater que : des communautés ont utilisé les fonds à travers des projets, d'autres ne l'ont pas fait. La base de données a pu être élaborée selon deux indicateurs : le nombre de projets par communauté et le ratio entre le montant de la totalité des projets/ l'enveloppe de dépenses disponible. L'analyse de la base de données nous permettra de constater qu'il est possible de classer les communautés selon trois grandes catégories : i) celles qui ont peu de projets et qui n'ont presque pas dépensé l'enveloppe attribuée; ii) celles qui ont peu de projets, mais qui ont dépensé la totalité (ou presque) de leur enveloppe; iii) celles qui ont beaucoup de projets et qui ont dépensé la totalité (ou presque) de leur enveloppe. Il apparaîtra alors intéressant de sélectionner au moins une communauté dans chacune des catégories. À cela pourra s'ajouter d'autres critères, par exemple : le nombre d'habitants de ces communautés, l'éloignement géographique de ces communautés par rapport à l'accès à des services publics, le PIB de ces communautés, la nature de leurs économies, etc. Plus le nombre de critères sera important, plus il sera difficile de faire des choix. Ainsi, il apparaît primordial que l'équipe

d'évaluation soit en mesure d'expliquer pourquoi certains critères sont retenus, et d'autres non, et ce choix se fait a minima avec la collaboration du client principal de l'évaluation. Certaines communautés peuvent avoir un poids politique important ou peuvent avoir posé des difficultés particulières pour la mise en œuvre de l'intervention : faudra-t-il les inclure au risque que l'évaluation accorde trop de place à leurs problématiques particulières, ou au contraire les exclure au risque d'avoir à gérer des conflits politiques par la suite? Il n'y a pas de réponse toute faite.

Parmi les échantillons non probabilistes, on peut aussi mentionner les échantillons dits « accidentels » (p. ex. interrogation d'individus au coin d'une rue), les échantillons de volontaires (faire appel à des volontaires pour l'étude), les échantillons en boule de neige (sélection de personnes qui sont mises en contact avec d'autres personnes), les échantillons par quotas (reproduction de la population à étudier en créant des sous-groupes sans sélection aléatoire des unités).

### **Ce qui caractérise les méthodes mixtes : de la fin d'une controverse à l'enjeu de l'intégration**

Les méthodes quantitatives et qualitatives se sont opposées par le passé, mais se sont également nourries mutuellement des critiques de l'une envers l'autre. Ces débats sont considérés comme résolus, mais ils semblent refaire surface dans le contexte de l'évaluation des effets (Ridde *et al.*, 2021). Privilégiant un certain pragmatisme, la majorité des évaluations incorporent désormais des stratégies mixtes de collecte et d'analyse des données qui allient des méthodes et des outils quantitatifs et qualitatifs; et impliquent donc des équipes multidisciplinaires (souvent qualifiées de « méthodes mixtes » pour simplifier). Ce type de stratégie permet d'optimiser les avantages et de parer aux limites des méthodes et des outils, chacun pris isolément, pour obtenir une plus grande compréhension de l'intervention. Lorsque des méthodes mixtes sont mises en œuvre, un important effort doit être réalisé quant au respect des normes propres à chaque méthode de nature quantitative et qualitative. Un effort supplémentaire, spécifique à une stratégie mixte, doit être réalisé pour s'assurer d'obtenir des constats valides, mais il n'y a pas encore un consensus concernant des critères de qualité spécifiques à une stratégie mixte. Le principal défi demeure *l'intégration* (la combinaison) des méthodes qualitatives et quantitatives entre elles puisque celles-ci relèvent de traditions différentes, notamment pour tirer des constats et des conclusions issus à fois de l'analyse et de l'interprétation des données quantitatives et qualitatives. Ainsi, l'évaluateur ou l'évaluatrice doit être en mesure de justifier et de donner un sens aux choix effectués pour chaque variable, indépendamment et dépendamment les unes des autres dans la matrice. Cela exclut la juxtaposition des méthodes et des outils sans réfléchir à la manière dont leurs techniques, leurs démarches, l'analyse et l'interprétation des données qu'ils permettent d'obtenir vont s'éclairer mutuellement (Pluye, 2019). On peut donc retenir comme critère de qualité d'une stratégie mixte : le degré de respect des critères de qualité de chaque méthode, la capacité de justifier les choix effectués dans la matrice, le degré d'intégration des méthodes, sa valeur ajoutée et ses limites.

Dans ce contexte, plusieurs auteurs ont décrit les différentes combinaisons possibles qui permettent une intégration réussie ainsi que les finalités associées à ces stratégies, notamment Creswell et Plano Clark (2018). La nature de ces combinaisons est au nombre de trois : devis à connexion de phases, devis comparatif et devis d'assimilation des données, mais l'ensemble des combinaisons possibles varient de sept à neuf selon les auteurs (Pluye, 2019). Ces combinaisons dépendent du séquençement choisi ou de l'agencement des méthodes et des outils ainsi que de la finalité qu'on leur donne.



Le devis à connexion de phases, aussi appelé devis séquentiel, devis à « développement séquentiel » ou « analyse mixte séquentielle », a pour postulat de départ que les interprétations des sciences liées aux méthodes qualitatives sont différentes de celles liées aux méthodes quantitatives (Bujold *et al.*, 2016). Ce devis consiste à mettre en œuvre une phase qualitative et une phase quantitative de façon séparée, mais une phase (ou plus) doit suggérer la complémentarité des données. Le processus d'intégration se trouve alors dans la connexion entre deux phases (Pluye, 2019). Les résultats d'une première phase doivent permettre d'élaborer l'échantillon ou les instruments de l'autre (Bamberger, 2012, p. 5). On peut ainsi construire une phase qualitative exploratoire, puis une phase quantitative pour favoriser la généralisation des constats; souhaiter obtenir de l'information complémentaire avec une phase qualitative à la suite d'une phase quantitative pour obtenir une ou des explications, ou encore approfondir les éléments d'une phase avec l'autre selon un « fil conducteur » (Bujold *et al.*, 2016; Nha Hong *et al.*, 2020).

Dans le devis à comparaison des résultats, les données qualitatives et quantitatives sont collectées en parallèle et comparées. Les méthodes de collecte et d'analyse peuvent être interconnectées à plusieurs moments. Une première stratégie consiste à comparer les résultats quantitatifs et qualitatifs obtenus de collecte et d'analyse de données indépendantes (Nha Hong *et al.*, 2020). Il s'agit du design ségrégué, où les données qualitatives et quantitatives peuvent être considérées comme complémentaires (Sandelowski *et al.*, 2006), c'est-à-dire qu'elles permettent de documenter différentes variables. La deuxième stratégie compare les résultats qualitatifs et quantitatifs obtenus de collecte et d'analyse de données interdépendantes (Nha Hong *et al.*, 2020), c'est-à-dire qu'elle porte sur l'étude de mêmes variables. Lorsque les résultats convergent, on parle de triangulation des données; lorsqu'ils divergent, cela oblige une nouvelle analyse ou un recadrage, mais il est également possible de centrer les analyses sur les divergences (contradictions, discordances ou dissonances) entre les résultats qualitatifs et quantitatifs (Bamberger, 2012; Bujold *et al.*, 2016; Nha Hong *et al.*, 2020).

Lorsque les données collectées à partir d'un outil sont transformées puis fusionnées avec celles d'un autre, on parle d'assimilation des données (Nha Hong *et al.*, 2020). Aussi appelée consolidation, fusion ou transformation des données, ce type d'intégration sert à consolider les données obtenues et à obtenir de nouvelles perspectives. Par exemple, il est possible d'assigner une valeur numérique aux données qualitatives sous une forme qui est compatible avec celle des données quantitatives (The Joanna Briggs Institute, 2014), ou encore de traiter des données quantitatives comme des données qualitatives (Nha Hong *et al.*, 2020).

Dans le contexte des évaluations, le choix de combinaison (appelé « devis » ou *design* en anglais) est tributaire des délais et des ressources imparties. La pratique semble plutôt orienter les évaluateurs et évaluatrices à adopter des devis où les outils et méthodes quantitatives et qualitatives sont mises en œuvre de manière parallèle dans une perspective de complémentarité ou de triangulation des données. Dans tous les cas, il sera important pour les évaluateurs et évaluatrices de bien connaître les avantages et les limites des outils de collecte de données. Cette connaissance et l'expérience permettront de juger de la pertinence et du réalisme du choix de la combinaison mixte possible.

## Les outils de collecte de données

Les outils de collecte de données sont les instruments d'application des méthodes quantitatives, qualitatives ou mixtes. Cette partie présente des outils « classiques » issus des sciences économiques et sociales et des outils participatifs, plutôt créés comme moyen d'atteindre des populations vulnérables par des organisations non gouvernementales (ONG) ou des organisations internationales.

### Les outils « classiques » de collecte de données

L'analyse de données issues de sources secondaires, le sondage, l'entrevue, le groupe de discussion et l'observation sont les outils les plus communément utilisés en évaluation. Le tableau suivant résume brièvement en quoi consiste chacun de ces outils, leurs avantages et leurs limites.

**Tableau 1 — Synthèse des avantages et des limites des outils classiques de collecte des données. À partir de l'expérience des auteurs et de Campenhoudt *et al.*, 2017, Durand, 2021, Martineau, 2021, Maltais *et al.*, 2021, Patton, 2014.<sup>6</sup>**

Outils	Qu'est-ce que c'est ?	Avantages	Limites	Commentaires
Analyse des données issues des sources secondaires	Analyse des données issues des sources secondaires, soit les données collectées dans une base de données spécifique à l'intervention, des données d'indicateurs plus larges, des données d'enquête ou de recensement d'un institut de statistique nationale.	<p>Les données sont facilement accessibles et peu coûteuses à obtenir.</p> <p>Le temps de collecte est relativement rapide.</p> <p>Elle permet d'optimiser le temps de collecte et de construire une stratégie subséquente.</p> <p>Les données disponibles permettront d'obtenir un portrait global de la situation ou d'avoir des informations complémentaires.</p> <p>Certaines bases de données comptent un grand nombre d'unités (plus d'un millier), ce qui permettra des analyses statistiques avancées intéressantes (Campenhoudt <i>et al.</i>, 2017).</p> <p>Elle devient de plus en plus pertinente dans le contexte de la transformation numérique où de plus en plus d'informations sont collectées à travers Internet (Campenhoudt <i>et al.</i>, 2017).</p>	<p>Elle ne répond pas directement aux questions d'évaluation (dépendamment des indicateurs choisis).</p> <p>Les informations sur la méthodologie ne sont pas toujours disponibles. Le contexte de collecte de ces données est souvent perdu.</p> <p>Les bases de données peuvent être incomplètes (données manquantes).</p>	Cette analyse, conjointement à l'analyse documentaire de l'intervention, devrait être réalisée en premier lieu. Elle permettra un gain de temps précieux sur les collectes et analyses subséquentes nécessaires. Comme mentionné dans le chapitre « La matrice d'évaluation », elle permettra aussi de sélectionner des indicateurs pour lesquels l'information est déjà disponible.

6. Nous ne ferons pas référence aux recensions des écrits, revues de littérature, analyses documentaires, etc., car ces outils devraient être mis en place de manière assez systématique dans la démarche d'évaluation pour construire un cadre théorique valide, une théorie du changement initiale, identifier les limites des données disponibles, identifier les intentions des acteurs et actrices, etc.

Outils	Qu'est-ce que c'est ?	Avantages	Limites	Commentaires
Sondage ou enquête	Questionnaire constitué majoritairement de questions fermées (choix de réponses préétablis) adressées à un échantillon d'une population ou à une population <sup>7</sup> .	<p>Seul instrument approprié pour effectuer la mesure d'une variable (c'est-à-dire collecter des observations sur un indicateur).</p> <p>Il permet de recueillir un nombre important d'informations.</p> <p>Il permet de faire plusieurs analyses statistiques avancées (notamment d'étudier le lien entre les variables), dépendamment de la taille de l'échantillon.</p> <p>Il permet la généralisation des constats, dépendamment de la technique d'échantillonnage (la représentativité de l'échantillon doit être garantie).</p> <p>Il permet de présenter les données de façon visuellement attrayante (tableau, graphique, etc.)</p> <p>Flexibilité et polyvalence (Durand, 2021).</p>	<p>Dépendamment du choix d'administration du questionnaire, la collecte peut être longue et coûteuse.</p> <p>Il n'est pas approprié pour les populations qui ne peuvent répondre aux questions telles que formulées (p. ex. populations analphabètes).</p> <p>Il est soumis au risque d'avoir un taux de réponse faible.</p> <p>Il permet seulement de collecter l'information demandée sans éléments de contexte.</p> <p>Le traitement des réponses pour des analyses statistiques avancées demande des compétences spécifiques.</p>	<p>Plusieurs modalités d'administration du questionnaire sont possibles (autoadministration, par téléphone, électronique, en personne, par la poste), chacune ayant leurs avantages et leurs limites.</p> <p>Pour être valide, un sondage doit réunir un ensemble de conditions préalables : la disponibilité des personnes sondées, la capacité de répondre (compréhension du sens des questions et des choix de réponses), la transmission et l'enregistrement fidèle de l'information (en lien avec l'administration du questionnaire et la sincérité des réponses obtenues) (Durand, 2021b, p. 329).</p> <p>La qualité du sondage est tributaire de la formulation des questions<sup>8</sup>, de la manière de l'administrer et de la gestion des refus de répondre.</p>

7. Durand (2021b) fait la différence entre sondage et enquête par la présence ou non d'un échantillon. Elle indique que le sondage correspond à « l'observation de réponses à des questions posées à un échantillon d'une population » (Durand, 2021b, p. 327), tandis que lorsque le questionnaire s'adresse à l'ensemble de la population, elle utilise les termes « enquête » ou « recensement ». De notre côté, étant donné l'usage très répandu du terme « enquête » pour parler de l'administration de questionnaires en face à face, y compris sur la base d'un échantillonnage; et le fait que, dans la pratique de l'évaluation, des questionnaires peuvent être adressés à des populations dans leur entièreté, car de taille réduite (l'échantillon représente alors le nombre de personnes ayant répondu au questionnaire), nous utilisons les termes « enquête » et « sondage » sans distinction. En revanche, le terme « recensement » fait clairement référence à l'administration d'un questionnaire à l'ensemble de la population (plutôt à l'échelle de la société).

8. Plusieurs critères doivent être respectés à cet effet : la précision et la neutralité des questions, l'usage d'échelles connues (surtout en évaluation), la longueur du questionnaire, l'ordre des questions, la possibilité de donner une porte de sortie (ne sais pas, ne veut pas répondre), etc. Voir à cet effet Durand, 2021b et Schaeffer *et al.*, 2003.

Outils	Qu'est-ce que c'est ?	Avantages	Limites	Commentaires
Entrevue ou entretien	Interaction verbale entre deux personnes, caractérisée par une série de questions ouvertes (sans choix de réponses préétablies) de la part d'une des personnes (intervieweur).	<p>Elle permet de collecter un nombre important d'informations contextualisées et en profondeur.</p> <p>Elle donne une grande liberté de réponses et permet de respecter l'intégrité de la personne interrogée (en termes de cadre de référence, de langage, de catégories mentales) (Campenhoudt <i>et al.</i>, 2017).</p> <p>La forme de l'entrevue peut être adaptée en fonction de la situation d'énonciation.</p>	<p>Le lien de confiance entre la personne répondante et l'intervieweur, qui permet d'obtenir des informations riches et sincères, est tributaire des savoirs, savoir-faire et savoir-être de ce dernier. Ces aspects s'acquièrent avec une bonne connaissance théorique des techniques d'entrevue, l'expérience, une capacité d'autocritique et de mise à distance de soi.</p> <p>Théoriquement, les informations obtenues en entrevue devraient être consignées au mot à mot (p. ex. transcription d'enregistrement), mais en évaluation, par manque de temps, des synthèses d'entrevue sont couramment utilisées, ce qui introduit un biais (mémoire sélective, perte d'information, perte d'éléments de contexte, etc.).</p> <p>Il est nécessaire d'effectuer un certain nombre d'entrevues pour parvenir à la saturation des données (c'est-à-dire que les informations recueillies d'une entrevue à l'autre ne changent plus vraiment).</p>	<p>Il existe plusieurs formes d'entrevue : structurée (un guide d'entrevue avec la formulation des questions préétablies), semi-structurée (un guide d'entrevue des thématiques sans formulation des questions) ou libre (comme une conversation).</p> <p>L'entrevue, ou l'entretien, peut être réalisée en face-à-face (y compris à distance avec la vidéoconférence) ou par téléphone.</p>

Outils	Qu'est-ce que c'est ?	Avantages	Limites	Commentaires
Groupe de discussion ou « entretien de groupe »	« Une discussion entre différentes personnes sur un sujet précis et qui est modérée par une personne qui en est responsable dans le but de générer des données » (Maltais <i>et al.</i> , 2021, p. 296).	Il est habituellement assez facile à mettre en œuvre. La formulation des questions est très souple, car il s'agit de favoriser une dynamique de groupe.	Il nécessite une bonne animation et une gestion des différents types de personnalités. Le succès est lié à la dynamique de groupe mise en place. Les avis divergents peuvent ne pas s'exprimer par conformisme avec l'opinion majoritaire du groupe. Il possède les mêmes enjeux de retranscription de l'information obtenue que pour l'entrevue, couplé à celui de rendre compte de la dynamique du groupe dans la prise de notes. La confidentialité n'est pas garantie. La présence d'une figure d'autorité peut reléguer des participants au silence ou au conformisme.	Le groupe de discussion est particulièrement intéressant lorsqu'il s'agit d'obtenir un consensus sur un sujet (débat), lorsqu'on souhaite étudier les interactions sociales, pour recueillir de l'information auprès de groupes particuliers ou de populations marginalisées, pour examiner des sujets délicats, pour explorer un phénomène vécu par plusieurs personnes (p. ex. les changements dus à une intervention) (Maltais <i>et al.</i> , 2021, p. 296).

Outils	Qu'est-ce que c'est ?	Avantages	Limites	Commentaires
Observation	« Une approche de recherche et outil de formation de l'information où la personne est le témoin – plus ou moins à distance – des comportements des individus et des pratiques au sein des groupes en séjournant sur les lieux mêmes où ils se déroulent » (Martineau, 2021, p. 234).	<p>Elle permet d'obtenir de l'information en temps réel, de façon directe, spontanée, voire inattendue (Campenhoudt <i>et al.</i>, 2017).</p> <p>Elle permet d'observer les comportements (plutôt que de les rapporter sous forme écrite ou par la parole) ainsi que les lieux (p. ex. l'existence réelle de livrables matériels de l'intervention) (Patton, 2014).</p> <p>Elle permet d'élargir les sources de données secondaires (p. ex. l'obtention de documents, la consultation de journaux locaux, etc.) (Patton, 2014).</p>	<p>Il est difficile de pouvoir planifier ce qui sera observé.</p> <p>Des techniques particulières de mise à distance de soi doivent être mises en place (notamment, la pratique d'un journal de bord).</p> <p>Le temps de préparation à l'entrée du terrain peut être long et nécessite une planification préalable.</p> <p>Les comportements des sujets peuvent changer par le fait de se savoir observé.</p> <p>Il est nécessaire d'être introduit sur le terrain par quelqu'un qui, en même temps, peut introduire un biais.</p> <p>L'interprétation de ce qui est observé doit faire l'objet d'une grande prudence, et nécessite souvent la participation des sujets observés (voir les outils de collecte de données participatifs qui peuvent alors enrichir les observations sur le terrain dans une perspective de collecte complémentaire) (Campenhoudt <i>et al.</i>, 2017; Patton, 2014; Martineau, 2021).</p>	Plusieurs éléments de la mise en œuvre de l'observation, qualifiée d'entrée sur le terrain, sont examinés dans le chapitre « L'entrée sur le terrain en évaluation : enjeux et pratiques des collectes des données ».

Dans une perspective de décolonisation de la recherche, des critiques ont émergé sur la justesse des interprétations issues de l'analyse des données obtenues avec ces outils; interprétations qui demeurent tributaires du cadre de référence des chercheurs et chercheuses ou des évaluateurs et évaluatrices (caractère stéréotypé, méconnaissance du contexte). Des méthodologies décolonisées, qui touchent l'ensemble de la démarche de recherche (ou dans le cadre de ce chapitre, la démarche d'évaluation), du choix des questions évaluatives à l'ensemble de la stratégie de collecte et d'analyse de données, émergent en Afrique et dans les milieux autochtones. L'une de leurs caractéristiques est la mise en place d'une collaboration avec les communautés auprès de qui l'information est collectée; les modes de partage de l'information ou du savoir sont ceux privilégiés par les communautés participantes (p. ex. le récit (*storytelling*), la transmission orale, etc.) (Maltais et Bourgeois, 2021). L'évaluation n'échappe pas à ces perspectives, d'autant plus que le contexte prend une place déterminante dans l'évaluation des besoins ou l'identification des problèmes, de la pertinence, des processus d'une intervention ou des changements qu'elle a pu engendrer. Par ailleurs, des méthodes féministes de collecte et d'analyse des données privilégient la création de sens à partir du point de vue de l'expérience partagée par des femmes ayant des caractéristiques communes (contextualisées selon une perspective intersectionnelle) et visent l'autonomisation (*empowerment*) des femmes participant à la collecte. La mise en œuvre concrète de la perspective «aucun laissé pour compte» dans la stratégie de collecte et d'analyse des données nécessite la mise en œuvre d'outils participatifs de collecte de données pour vraiment faire entendre les voix des populations marginalisées, vulnérables, «sans pouvoir».

### **Les outils participatifs de collecte de données**

Plusieurs des outils présentés ici sont nés du contexte des projets de la coopération internationale ou de situations de recherche particulière. Plusieurs principes d'application de ces outils et techniques de mise en œuvre découlent généralement des outils qualitatifs «classiques».

#### ***La carte communautaire des performances***

La carte communautaire des performances (CCP) est un outil participatif d'analyse des services. La partie qui demande un service, les usagers du service, et celle qui fournit ce service, le prestataire de service, sont interrogées séparément, sous forme d'entretien de groupe, pour analyser le service. Une réunion qui implique les deux parties est ensuite mise en œuvre pour trouver des solutions acceptées par tous aux problèmes posés (p. ex. de qualité, d'accessibilité ou d'utilisation du service) et pour déboucher sur un plan d'action. Facile d'utilisation, cet outil est mis en œuvre au niveau local, particulièrement dans les secteurs de la santé, de l'éducation, de l'agriculture et de la gestion de l'eau, mais il peut être adapté à n'importe quel secteur (Kiracho *et al.*, 2020). Il est idéal pour obtenir de l'information immédiate, et est particulièrement apprécié dans les communautés rurales plus marginalisées (Kiracho *et al.*, 2020; Mwanza et Ghambi, 2011). Il contribue à augmenter la participation, la redevabilité et la transparence entre usagers, prestataires de services et décideurs ou décideuses (Escher, 2018). Il permet de clarifier les responsabilités de chacun et de dénoncer des abus. Toutefois, il demande beaucoup de temps, l'obtention de résultats concrets demeure donc complexe. Il peut aussi conduire à des conflits ou à des attentes exagérées de la part des communautés s'il est mal utilisé ou mal compris (Care, 2015).



## Encadré 2 – Évaluation de la qualité des services de santé et de leur accessibilité par Save the Children en 2019

Pour améliorer l'accès à des services de santé de qualité dans le cadre de la mise en œuvre d'un programme en République démocratique du Congo, une équipe d'évaluation et d'amélioration de la qualité a collaboré avec des prestataires de services et des représentants des communautés les plus marginalisées des zones de santé ciblées dans la province du Kasai Oriental. Il s'agissait d'identifier les problèmes de services et de développer des indicateurs de qualité pour obtenir une carte de scores. L'objectif ultime de cette évaluation étant de combler le fossé entre les perceptions de la communauté et celles des prestataires sur la qualité des services, les défis majeurs ont porté sur l'engagement des acteurs communautaires tout au long de l'évaluation ainsi que sur les enjeux éthiques de la facilitation des échanges : comment faire participer les femmes de la communauté pour qu'elles se prononcent sur les questions de santé sexuelle et reproductive dans un contexte où la présence des pairs ou le profil des animateurs et animatrices dans les entretiens de groupe peuvent être intimidants ?

### La méthode photovoix (PhotoVoice)

Il peut être difficile pour les évaluateurs d'observer certains des comportements les plus importants d'une communauté, en particulier lorsque cette communauté se méfie d'étrangers ou lorsqu'elle opère sans langage écrit. Dans ces situations, l'observation peut aboutir à une vision biaisée et inexacte de la communauté, ce qui peut alors entraîner une mauvaise compréhension de ses besoins, de ses atouts et de ses valeurs. Selon la traduction de Rey *et al.*, 2019, de Wang et Burris, 1997, p. 369 :

La photovoix est un processus par lequel les individus peuvent identifier, représenter et améliorer leur communauté à travers l'utilisation d'une technique photographique spécifique, en confiant des caméras à des personnes pour leur permettre d'agir comme « enregistreurs » et catalyseurs potentiels du changement dans leurs propres communautés.

Après s'être mis d'accord sur l'objet des photographies avec la communauté, grâce aux caméras fournies par l'équipe d'évaluation, des membres de la communauté sont invités à capturer des représentations visuelles de leurs expériences par rapport à une intervention ou des problèmes particuliers<sup>9</sup>. Par la suite, lors d'une phase narrative, ces photographies sont discutées, décodées et interprétées par le biais d'entretiens individuels et d'entretiens de groupes avec les membres de la communauté. Cette méthode permet aux évaluateurs, dans un esprit de dialogue avec les communautés, d'avoir un aperçu des pratiques invisibles, de capturer le ressenti des personnes vis-à-vis de leur environnement, des changements, des besoins ou des problèmes qui ont du sens pour eux. La plupart du temps, elle est utilisée avec des groupes ou des individus vivant dans la marginalité, dans le but de comprendre leurs besoins (Molloy, 2007; Plunkett *et al.*, 2012; Strack *et al.*, 2004) et elle a particulièrement été utilisée afin d'étudier l'évolution de l'autonomisation de communautés plus vulnérables (Budig *et al.*, 2018).

9. Les photographies peuvent être capturées dans le cadre des mouvements effectués durant la journée par la personne qui est munie de la caméra, d'où le terme « *photowalk* » en anglais, littéralement « photo en mouvement », qui fait partie de la pratique de cette méthode.



Outre le caractère innovant et immersif de la méthode, parmi ses avantages, on peut mentionner la participation communautaire, une communication adaptée à l'audience, la créativité et la solidarité communautaire. Parmi les limites, la méthode est coûteuse à mettre en place, prend du temps (Nykiforuk *et al.*, 2011) et soulève des enjeux éthiques en lien avec le caractère intrusif de la méthode (Asaba *et al.*, 2014), l'obtention du consentement pour la prise de photos, la distribution des caméras, la perception de la communauté sur les personnes sélectionnées pour porter les caméras, le désir de plaire (prendre des photos qui feront plaisir à l'équipe d'évaluation, mais qui ne reflètent pas nécessairement l'expérience ou le vécu) (Evans-Agnew et Rosemberg, 2016), la réaction négative de la communauté face aux images prises, comme preuves de leur pauvreté et de violence et de dysfonctionnements qui peuvent entrer en contradiction avec les messages ou l'image que la communauté souhaite véhiculer auprès des étrangers (Milne et Muir, 2019).

### **Encadré 3 – Mise en œuvre de la méthode photovoix pour une évaluation de référence sur un programme de malnutrition en 2019**

L'Université de Kinshasa a utilisé la méthode photovoix en vue de saisir les facteurs pouvant expliquer la malnutrition dans les communautés visitées dans le Kasai Oriental et dans la province du Lomami en République démocratique du Congo. Les participants choisis sur un échantillon de 950 ménages dans les deux provinces se sont prononcés grâce aux photos prises sur les comportements alimentaires et les stratégies de survie. Une formation portant sur l'utilisation de tablettes, la prise et la sélection d'images, ainsi que sur la manière de réaliser la narration de la signification d'images lors d'entretiens de groupes et l'autoréflexion sur lesdites images obtenues a été conduite par le département de santé de l'Université de Kinshasa. Les membres de la communauté sélectionnés de manière raisonnée (échantillon non probabiliste), en collaboration avec les relais communautaires, ont eu 20 minutes chacun pour aller prendre des photos, les enregistrer et revenir au site de discussion de groupes. Chaque personne ayant pris une photo a été invitée à présenter sa photo au groupe et à la décrire. Douze superviseurs ont pu apporter un accompagnement technique tout au long du processus pour s'assurer que les tablettes sont bien utilisées et que l'accent est véritablement mis sur les facteurs pouvant expliquer la malnutrition dans les communautés visitées. Le temps de mise en œuvre de la collecte a représenté un défi majeur, car les membres de la communauté n'étaient pas familiers avec la technologie et le processus de collecte, et des photos de qualité insuffisante ont dû être reprises.

### **Vote par points**

Cet outil permet d'identifier et de prioriser les principaux problèmes auxquels les répondants sont confrontés par un système de vote par points. Lors de séances d'une durée de 45 minutes, chaque personne obtient trois votes pour identifier les questions ou les problèmes considérés comme prioritaires. Les trois problèmes qui ont obtenu le plus de votes sont discutés dans le cadre de groupes de discussion. Cet outil est idéal pour de grands groupes dans lesquels un grand nombre d'idées sont identifiées, et il peut s'adresser autant aux adultes qu'aux enfants. Il évite le formalisme d'un questionnaire de sondage. Cependant, le caractère public des votes peut exercer une influence sur les

participants (Steven, 2020). Il comporte également les limites d'un entretien de groupe, notamment le désir de conformisme avec l'opinion majoritaire et la gestion de la dynamique de groupe qui demeure ainsi tributaire des qualités d'animation du modérateur ou de la modératrice.

#### **Encadré 4 – Exemple de mise en œuvre d'un vote par points en République démocratique du Congo par Save the Children**

Dans le cadre de la collecte de données pour l'évaluation des besoins des enfants, des jeunes et des adultes en 2019, le vote par points a été utilisé auprès de 2437 enfants et 234 adultes à travers différentes consultations. Dans la région du Kivu, 58 groupes de discussion ont été organisés, avec un total de 596 participants, soit une moyenne de 10,2 participants par groupe de discussion. Outre les organisations locales, les données ont été collectées avec le soutien de plusieurs institutions y compris la Division des affaires sociales (DIVAS), la Division provinciale des actions humanitaires (DIVAH), la Division provinciale de la santé (DPSP) et la Division provinciale de l'éducation (EPSP). Dans toutes les consultations, un remue-méninges participatif sur les problèmes rencontrés par les enfants, les jeunes et les adultes suivi d'un vote par points ont été réalisés pour les amener à classer leurs problèmes par ordre de priorité. Il s'est ensuivi des discussions ouvertes sur les choix effectués. Sur le plan éthique, l'équipe est restée intentionnellement souple par rapport à l'activité, et en a limité la durée pour respecter les activités quotidiennes des enfants et des membres de la communauté (agriculteurs, pêcheurs, etc.). L'équipe a reçu le quitus de la Commission nationale d'éthique de Kinshasa, et tous les formulaires de consentement dédiés aussi bien aux prestataires qu'aux enfants et aux jeunes ont été dûment signés par ceux-ci. L'utilisation de l'outil a permis de susciter l'enthousiasme des répondants en leur donnant le contrôle et en évitant le format conventionnel d'un questionnaire. Afin de retenir l'attention des enfants, des jeunes et des adultes (dont des personnes non instruites), l'équipe d'évaluation a utilisé une approche assez ludique par l'utilisation d'images pour matérialiser les problèmes identifiés. L'appréciation de l'outil et l'obtention de réponses franches ont été une réussite pour cette collecte.

#### ***Cartographie corporelle ou cartographie du corps***

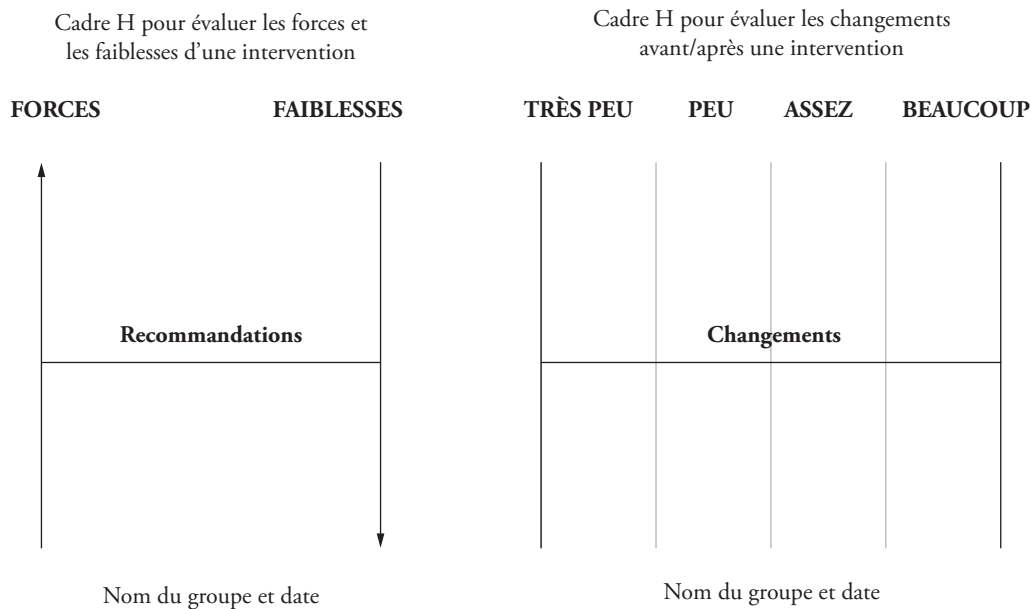
La cartographie corporelle est une méthode où, à l'aide d'une représentation visuelle du corps (dessin), les personnes interrogées sont encouragées à penser aux différentes parties du corps pour explorer et noter les changements vécus après leur participation à une intervention (ou en recherche en lien avec un élément déclencheur de changements). À partir d'un processus artistique, la cartographie du corps permet de collecter de nombreux types d'informations en les reliant au corps (p. ex. la tête peut faire référence aux changements au niveau des connaissances ou de la pensée, le cœur en lien avec les émotions, les yeux en lien avec la façon de se voir ou de voir son entourage, etc.) (Lys *et al.*, 2018; Save the Children, 2014).

La première utilisation documentée de la cartographie corporelle remonte à 1987, dans le cadre d'une étude sur la fertilité, en Jamaïque et en Angleterre. Dans les années 2000, celle-ci s'utilise en contexte africain, notamment auprès de femmes vivant avec le sida ou le VIH (Vasquez, 2004). Elle a aussi été utilisée à des fins politiques, avec le groupe de femmes de Bambini, un projet qui visait à rendre plus de médicaments antiviraux disponibles. L'outil permet d'accéder à des perspectives uniques, notamment parce qu'il est accessible à tous. Un tracé de corps est réalisé autour du corps du participant ou du chercheur, mais qui réfère toutefois au participant. Sous la forme de réflexions ou d'entretiens des symboles et/ou des slogans personnels sont dessinés dans le tracé, le tout générant un narratif personnel. Finalement, la cartographie est alimentée par la réflexion des autres et est donc à la vue de tous. La cartographie a plusieurs finalités. Tout d'abord, elle permet de partager différents narratifs, dans un renversement des dynamiques de pouvoir, puisque le participant fait des choix artistiques. Elle peut aussi servir de forme de plaidoyer politique, lorsqu'on donne la parole à des communautés marginalisées. La diffusion et la présentation des résultats dans des endroits stratégiques peut offrir au grand public de nouvelles perspectives. Elle peut aussi servir d'outil éducatif, notamment dans un objectif de déstigmatisation. Elle permet également de mieux comprendre l'expérience incarnée, par la démonstration des expériences sensorielles et émotionnelles de participants en lien avec certains lieux ou espaces. Elle est aussi idéale pour l'évaluation avec des adolescents ou des enfants, afin qu'ils puissent formuler leurs narratifs de leur propre façon, sans certaines conceptions adultes (De Jager *et al.*, 2016).

Parmi les leçons apprises issues de la pratique, il est essentiel de conserver les données de référence en gardant soigneusement la carte du corps « avant » l'intervention, afin de pouvoir la ressortir et l'utiliser au moment d'une évaluation *ex post*. Lors d'une expérience au Burkina Faso, il a été possible de mettre à jour la carte au fur et mesure des progrès réalisés, ce qui a permis de mieux apprécier les changements et de porter un jugement plus objectif lors des évaluations planifiées. Il est aussi crucial de prendre en compte les caractéristiques de genre, de handicap et de participation pour anticiper sur les enjeux logistiques et éthiques et mettre en œuvre l'outil dans un contexte sécuritaire (notamment pour les sujets délicats comme les violences sexuelles ou les maladies) avec des facilitatrices et facilitateurs expérimentés.

### *Cadre H*

Le cadre H correspond à un entretien de groupe à partir du dessin d'un H où la ligne horizontale touchant le côté gauche du H correspond au point 0, et le côté droit au point 10. Les participants et participantes se mettront d'accord avec le facilitateur ou la facilitatrice sur la thématique abordée; ils devront placer leur position sur la ligne horizontale du H en fonction de leur accord ou non avec une série d'énoncés en lien avec la thématique. Les résultats sont ensuite discutés en groupe. En évaluation, il est particulièrement utile pour identifier les forces (à gauche du H) et les faiblesses (à droite du H) de l'intervention, ou encore pour identifier les changements à l'issue d'une intervention (peu ou pas du tout de changements à gauche du H; changements significatifs à droite du H), telle que l'illustre la figure suivante.



**Figure 1**

Exemple de cadre H pour une évaluation

**Encadré 5 – Exemple de mise en œuvre du cadre H par le ministère finlandais des affaires étrangères, Save the Children et le Mouvement Africain des Enfants et Jeunes Travailleurs**

Dans le cadre d'une évaluation à mi-parcours d'un programme régional, conduite en 2016 et couvrant huit pays d'Afrique de l'Ouest et impliquant le Mouvement Africain des Enfants et Jeunes travailleurs, l'utilisation du cadre H a permis de cerner les aspects contribuant à la réduction des risques de catastrophes en analysant le développement et l'application systématique des politiques, des stratégies et des pratiques, ainsi que les éléments programmatiques contribuant positivement et négativement à la redevabilité et à la transparence. À partir de l'analyse des forces et des faiblesses du programme, les jeunes des pays concernés et les responsables nationaux ont proposé des recommandations et des plans d'action pertinents pour chaque coordination nationale. Le contexte de mise en œuvre de cet outil a permis d'obtenir des réponses sensibles au genre. Le principal défi rencontré concernait la possibilité de retranscrire et de synthétiser de façon concise et précise les données recueillies dans le H par les jeunes.

Il existe bien d'autres outils participatifs qui peuvent faire partie du bagage de l'évaluateur ou de l'évaluatrice (p. ex. l'animal vivant, la montgolfière, etc.). Les outils participatifs présentent l'avantage de promouvoir l'engagement et la redevabilité envers les parties prenantes. Ils mettent en relief les normes, les comportements et les attitudes clés à prendre en compte pour les évaluateurs, et ils leur

permettent ainsi d'être des auditeurs actifs. Ces outils partent du principe que la façon dont chaque personne conçoit sa situation est tout aussi valable que la conception d'une autre personne. L'efficacité de l'utilisation des outils participatifs réside ainsi dans le processus, plutôt que dans les outils utilisés. Ils peuvent être aussi utiles avec des enfants qu'avec des adultes dans divers contextes, car la collecte des données ne repose pas sur l'alphabétisme des personnes, mais met un plus grand accent sur le pouvoir des impressions visuelles et sur la représentation active des idées.

Le tableau suivant présente les avantages et les limites des outils participatifs de collecte de données.

**Tableau 2 — Avantages et inconvénients des outils participatifs de collecte**

Avantages	Inconvénients
<p>Ils permettent de rejoindre des populations vulnérables, difficiles d'accès ou qui sont peu consultées, selon un processus d'autonomisation.</p> <p>Ils permettent l'obtention d'information en temps réel, qui n'aurait pu être obtenue autrement.</p> <p>Les constats ont du sens pour les participants, ce qui renforce l'obtention de constats utiles pour l'amélioration d'une intervention.</p> <p>Ils permettent des applications pratiques de plaidoyer.</p> <p>Il s'agit généralement d'outils pédagogiques, flexibles et ludiques.</p>	<p>Ils sont souvent longs à mettre en place, coûteux et nécessitent une logistique particulière.</p> <p>Ils sont tributaires du savoir, savoir-faire et savoir-être des personnes facilitatrices (en lien avec les entretiens et entretiens de groupe). Ils nécessitent des compétences particulières (p. ex. la cartographie du corps qui peut être chargée émotionnellement pour des sujets difficiles).</p> <p>Le risque d'une mauvaise utilisation peut porter préjudice aux populations (attention au principe « ne pas nuire ») et il faut rester attentif aux enjeux éthiques et au contexte sécuritaire de la mise en œuvre.</p>

## Conclusion

Nous avons présenté un bref panorama des outils de collecte de données, tant classiques que participatifs, utilisés en évaluation. Les débats sur la supériorité d'une stratégie quantitative par rapport à une stratégie qualitative sont considérés comme résolus par la communauté des chercheurs et des évaluateurs au profit des stratégies mixtes. L'évaluatrice ou l'évaluateur devrait bien connaître les avantages et les limites de chaque outil. Il devrait également être en mesure de justifier son choix de stratégie, expliquer les combinaisons effectuées (dans une perspective d'intégration) et pour quelles finalités. Ainsi, même si la démarche évaluative peut être globalement qualifiée de plus pragmatique, il n'en demeure pas moins que les évaluateurs et évaluateuses ont tout intérêt à enrichir leurs connaissances sur les fondements de chacune de ces démarches (paradigmes sous-jacents, démarches et techniques spécifiques à chaque outil, etc.) afin d'être en mesure d'identifier les limites méthodologiques de leur stratégie, qui ne sera jamais parfaite. Enfin, plusieurs espaces inédits de collecte et d'analyse de données s'offrent aux évaluateuses et évaluateurs, ceux de l'analyse des mégadonnées et de l'étude des réseaux sociaux. Les réseaux sociaux impliquent de nouvelles façons d'interagir, sont des lieux de coconstruction de comportements bénéfiques ou non, de critiques, d'espaces de mobilisation, etc. qui chamboulent la compréhension des comportements sociaux et de l'acceptabilité des interventions publiques. Les évaluateurs et évaluateuses ont donc tout intérêt à les prendre en compte dans la construction de leur démarche évaluative.

## Pour aller plus loin

Bourgeois, I. (2021). *Recherche sociale. De la problématique à la collecte des données* (7<sup>e</sup> édition). Presses de l'Université du Québec.

Cet ouvrage demeure une bonne entrée en matière, non seulement pour identifier les différents outils et méthodes classiques de la recherche en sciences sociales, mais pour mieux comprendre d'où vient la démarche systématique de collecte et d'analyse de données utile à l'évaluatrice ou évaluateur rigoureux.

Bamberger, M. (2012). *Introduction aux méthodes mixtes dans l'évaluation d'impact* (note n° 3). InterAction et The Rockefeller Foundation. <https://greval.ch/wp-content/uploads/2016/06/3-Mixed-Methods-in-Impact-Evaluation-FRENCH.pdf>

Cette publication est intéressante pour faire le tour des méthodes mixtes dans le cadre de l'évaluation des effets. Elle donne des illustrations concrètes des différentes stratégies possibles, y compris à des fins d'échantillonnage.

Enfin, nous recommandons des lectures pour aller plus loin sur la découverte d'outils participatifs :

CARE. (2015). *La carte communautaire de performances (CCP) : guide générique pour la mise en œuvre de la CCP de CARE pour améliorer la qualité des services*. [http://governance.careinternationalwikis.org/\\_media/ccp-boite-a-outils\\_care\\_2015.pdf](http://governance.careinternationalwikis.org/_media/ccp-boite-a-outils_care_2015.pdf)

Rey, L., Affodégon, W., Viens, I., Fathallah, H. et Arauz, M. J. (2019). La méthode photovoix. Dans V. Ridde et C. Dagenais (dir.), *Évaluation des interventions de santé mondiale* (p. 83-123). Éditions Sciences et bien commun et IRD Éditions. <https://scienceetbiencommun.pressbooks.pub/evalsantemondiale/chapter/photovoix/>

Save the Children. (2014). *Trousse à outils du suivi et de l'évaluation de la participation des enfants. Livret 5 : Outils de suivi et d'évaluation de la participation des enfants*. [https://resource-centre-uploads.s3.amazonaws.com/uploads/me\\_toolkit\\_booklet\\_5\\_french.pdf](https://resource-centre-uploads.s3.amazonaws.com/uploads/me_toolkit_booklet_5_french.pdf)

## Références bibliographiques

Anstett, E. (2016). *La mesure des effets. Programme international de formation en évaluation du développement* [matériel pédagogique, module 5]. École nationale d'administration publique.

Asaba, E., Laliberte Rudman, D. et Mondaca, M. (2014). Visual methodologies : photovoice in focus. Dans S. Nayar et M. Stanley (dir.), *Qualitative Research Methodologies for Occupational Science and Therapy*. Routledge.

Bamberger, M. (2012). *Introduction aux méthodes mixtes dans l'évaluation d'impact* (note n° 3). InterAction et The Rockefeller Foundation. <https://greval.ch/wp-content/uploads/2016/06/3-Mixed-Methods-in-Impact-Evaluation-FRENCH.pdf>

Beaud, J.-P. (2021). L'échantillonnage. Dans I. Bourgeois (dir.), *De la problématique à la collecte des données* (7<sup>e</sup> édition). Presses de l'Université du Québec.

Brett-MacLean, P. (2009). Body mapping : embodying the self living with HIV/AIDS. *Canadian Medical Association Journal*, 180(7), 740-741. <https://doi.org/10.1503/cmaj.090357>



- Budig, K., Diez, J., Conde, P., Sastre, M., Hernán, M. et Franco, M. (2018). Photovoice and empowerment : evaluating the transformative potential of a participatory action research project. *BMC Public Health*, 18(432). <https://doi.org/10.1186/s12889-018-5335-7>
- Bujold, M., Nha Hong, Q., Ridde, V., Bourque, C. J., Dogba, M. J., Vedel, I. et Pluve, P. (2016). Oser les défis des méthodes mixtes en sciences sociales et sciences de la santé. *Cahiers scientifiques de l'ACFAS*, 117.
- Campehouth, L. V., Marquet, J. et Raymond, Q. (2017). *Manuel de recherche en sciences sociales* (5<sup>e</sup> édition). Dunod.
- CARE. (2015). *La carte communautaire de performances (CCP) : guide générique pour la mise en œuvre de la CCP de CARE pour améliorer la qualité des services*. [http://governance.careinternationalwikis.org/\\_media/ccp-boite-a-outils\\_care\\_2015.pdf](http://governance.careinternationalwikis.org/_media/ccp-boite-a-outils_care_2015.pdf)
- Creswell, W. J. et Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research*. Sage Publications.
- Dagenais, C. et Ridde, V. (2012). *Approches et pratiques en évaluation de programmes*. Presses de l'Université de Montréal.
- De Jager, A., Tewson, A., Ludlow, B. et Boydell, K. M. (2016). Embodied Ways of Storying the Self : A Systematic Review of Body-Mapping. *Forum Qualitative Sozialforschung*, 17(2), <https://www.qualitative-research.net/index.php/fqs/article/download/2526/3986?inline=1#g2>
- Delahais, T., Devaux-Spatarakis, A., Revillard, A. et Ridde, V. (2021). Introduction générale. Dans T. Delahais, A. Devaux-Spatarakis, A. Revillard et V. Ridde (dir.), *Évaluation. Fondements, controverses, perspectives*. Éditions science et bien commun.
- Durand, C. (2021a). La mesure. Dans I. Bourgeois (dir.), *De la problématique à la collecte des données* (7<sup>e</sup> édition). Presses de l'Université du Québec.
- Durand, C. (2021b). Le sondage. Dans I. Bourgeois (dir.), *De la problématique à la collecte des données* (7<sup>e</sup> édition). Presses de l'Université du Québec.
- Escher, J. (2018). *Community Score Cards – a powerful tool to improve public services*. Civil Society Academy. <https://www.civilsocietyacademy.org/post/community-score-cards-a-powerful-tool>
- Evans-Agnew, R. A. et Rosember, M.-A. S. (2016). Questioning Photovoice Research : Whose Voice? *Qualitative Health Research*, 26(8), 1019-1030. <https://journals.sagepub.com/doi/abs/10.1177/1049732315624223>
- Fortin, M.-F. et Gagnon, J. (2016). *Fondements et étapes du processus de recherche. Méthodes quantitatives et qualitatives* (3<sup>e</sup> édition). Chenelière Éducation.
- Gauthier, B. (2021). La structure de la preuve. Dans I. Bourgeois (dir.), *De la problématique à la collecte des données* (7<sup>e</sup> édition). Presses de l'Université du Québec.
- Kiracho, E. E., Namuhani, N., Apolot, R. R., Aanyu, C., Mutebi, A., Tetui, M., Kiwanuka, S. N., Ayen, F. A., Mwesige, D., Bumbha, A., Paina, L. et Peters, D. H. (2020). Influence of community scorecards on maternal and newborn health service delivery and utilization. *International Journal for Equity in Health*, 19(145). <https://doi.org/10.1186/s12939-020-01184-6>
- Lincoln, Y. S. et Guba, E. G. (1985). *Naturalistic Inquiry*. Sage Publications.

- Lys, C., Gesink, D., Strike, C. et Larkin, J. (2018). Body Mapping as a Youth Sexual Health Intervention and Data Collection Tool. *Qualitative Health Research*, 28(7), 1185-1198. <https://doi.org/10.1177/1049732317750862>
- Maltais, S. et Bourgeois, I. (2021). Les enjeux contemporains de la recherche sociale. Dans I. Bourgeois (dir.), *De la problématique à la collecte des données* (7<sup>e</sup> édition). Presses de l'Université du Québec.
- Maltais, S., Auclair, I. et Brière, S. (2021). Les entretiens de groupe. Dans I. Bourgeois (dir.), *De la problématique à la collecte des données* (7<sup>e</sup> édition). Presses de l'Université du Québec.
- Martineau, S. (2021). L'observation directe. Dans I. Bourgeois (dir.), *De la problématique à la collecte des données* (7<sup>e</sup> édition). Presses de l'Université du Québec.
- Milne, E. J. et Muir, R. (2019). Photovoice : A Critical Introduction. Dans L. Pauwels et D. Mannay (dir.), *The SAGE Handbook of Visual Research Methods*. SAGE Publications.
- Molloy, J. K. (2007). Photovoice as a Tool for Social Justice Workers. *Journal of Progressive Human Services*, 2(18), 39-55. [https://doi.org/10.1300/J059v18n02\\_04](https://doi.org/10.1300/J059v18n02_04)
- Mwanza, J. et Ghambi, N. (2011). The community scorecard process : methodology, use, successes, challenges and opportunities. Dans The International Institute for Environment and Development (dir.), *Young citizens : youth and participatory governance in Africa* (p. 187-194). <https://resourcecentre.savethechildren.net/sites/default/files/documents/5504.pdf#page=189>
- Nha Hong, Q., Rees, R., Sutcliffe, K. et Thomas, J. (2020). Variations of mixed methods reviews approaches : A case study. *Research Synthesis Methods*, 11(6), 795-811. <https://doi.org/10.1002/jrsm.1437>
- Nykiforuk, C. I. J., Vallianatos, H., Nieuwendyk, L. M. (2011). Photovoice as a Method for Revealing Community Perceptions of the Built and Social Environment. *The International Journal of Qualitative Methods*, 10(2), 103-124. <https://doi.org/10.1177/160940691101000201>
- Patton, M. Q. (2014). *Qualitative Research and Evaluation Methods* (4<sup>e</sup> édition). SAGE Publications.
- Plunkett, R., Leipert, B. D. et Ray, S. L. (2012). Unspoken phenomena : using the photovoice method to enrich phenomenological inquiry. *Nursing Inquiry*, 20(2), 1-9. <https://doi.org/10.1111/j.1440-1800.2012.00594.x>
- Pluye, P. (2019). L'intégration en méthodes mixtes. Dans V. Ridde, et C. Dagenais (dir.), *Évaluation des interventions de santé mondiale* (p. 187-212). Éditions de l'Institut de recherche pour le développement. [10.4000/books.irdeditions.33393](https://doi.org/10.4000/books.irdeditions.33393)
- Rey, L., Affodégon, W., Viens, I., Fathallah, H. et Arauz, M. J. (2019). La méthode photovoix. Dans V. Ridde et C. Dagenais (dir.), *Évaluation des interventions de santé mondiale* (p. 83-123). Éditions Sciences et bien commun et IRD Éditions. <https://scienceetbiencommun.pressbooks.pub/evalsantemondiale/chapter/photovoix/>
- Ridde, V., Delahais, T., Devaux-Spatarakis, A. et Revillard, A. (2021). Introduction : la pluralité des approches paradigmatiques. Dans T. Delahais, A. Devaux-Spatarakis, A. Revillard et V. Ridde (dir.), *Évaluation. Fondements, controverses, perspectives*. Éditions science et bien commun.
- Rossi, P. H., Lipsey, M. W. et Freeman, H. E. (2004). *Evaluation : A systematic approach*. 7<sup>e</sup> édition. Thousand Oaks : Sage Publications.



- Roy, S. N. (2021). L'étude de cas. Dans I. Bourgeois (dir.), *De la problématique à la collecte des données* (7<sup>e</sup> édition). Presses de l'Université du Québec.
- Sandelowski, M., Voils, C. I. et Barroso, J. (2006). Defining and designing Mixed Research Synthesis Studies. *Research in the Schools*, 1(13), 29-40.
- Save the Children. (2014). *Trousse à outils du suivi et de l'évaluation de la participation des enfants. Livret 5 : Outils de suivi et d'évaluation de la participation des enfants*. [https://resource-centre-uploads.s3.amazonaws.com/uploads/me\\_toolkit\\_booklet\\_5\\_french.pdf](https://resource-centre-uploads.s3.amazonaws.com/uploads/me_toolkit_booklet_5_french.pdf)
- Schaeffer, N. C et Presser, S. (2003). The Science of Asking Questions. *Annual Review of Sociology*, 29, 65-88.
- Stevens, K. (2020). *Dotmocracy*. Better Evaluation. <https://www.betterevaluation.org/en/evaluation-options/Dotmocracy>
- Strack, R. W., Magill, C. et McDonagh, K. (2004). Empowering Youth through Photovoice. *Health Promotion Practice*, 5(1), 49-58. <https://doi.org/10.1177/1524839903258015>
- The Joanna Briggs Institute. (2014). *Methodology for JBI Mixed Methods Systematic Reviews*. University of Adelaide.
- Unité de l'impact et de l'innovation. (2019). *Modèles de mesure des impacts : Guide des méthodes de mesure des impacts*. Gouvernement du Canada. <https://www.canada.ca/fr/centre-innovation/services/rapports-ressources/modeles-mesure-impacts.html>
- Vasquez, G. (2004). *Body perceptions of HIV and AIDS : the memory Box Project*. Centre for Social Science Research, University of Cape Town. [https://open.uct.ac.za/bitstream/handle/11427/19177/Vasquez\\_Body\\_perceptions\\_2004.pdf?sequence=1&isAllowed=y](https://open.uct.ac.za/bitstream/handle/11427/19177/Vasquez_Body_perceptions_2004.pdf?sequence=1&isAllowed=y)
- Wang, C. et Burris, M. A. (1997). Photovoice : Concept, methodology, and use for participatory needs assessment. *Health education & behavior*, 24(3), 369-387.



# Les méthodes qualitatives et quantitatives pour évaluer les effets



*Eva Anstett  
Aref Ben Abdallah*

## Introduction

La question fondamentale au cœur de l'évaluation des effets est la suivante : comment pouvons-nous nous assurer que des changements observés, en lien avec les variables d'effets d'une intervention, ont bien été produits par l'intervention (Hempel et Fiala, 2011)? Établir ce lien entre une intervention et des changements est un problème théorique, qui se pose donc pour les questions d'évaluation dites explicatives ou de causalité, et qui n'existe pas pour d'autres types de questions d'évaluation, et pour lesquelles on peut donc fournir une réponse plus directement.

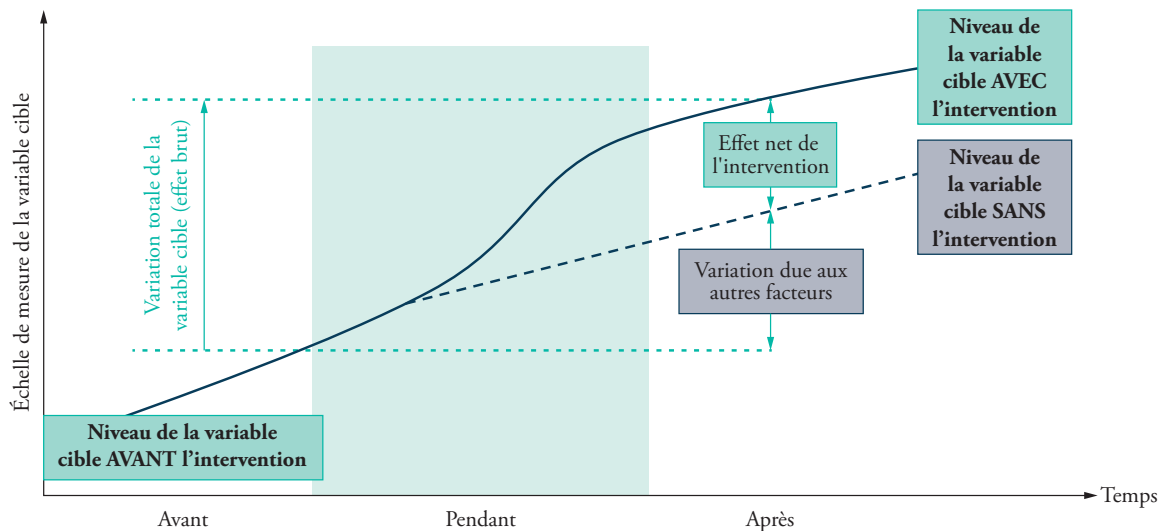
Ce défi théorique fait référence à la notion d'inférence causale, c'est-à-dire l'élaboration d'une démonstration empirique de la mesure dans laquelle une intervention donnée — et uniquement cette intervention — a contribué à changer une situation. Pour surmonter ce défi, et tenter d'établir un lien de causalité entre intervention et résultats, les équipes d'évaluation utilisent une diversité de méthodes de mesures, de calculs et de collecte de données sur le terrain, ce qui a fait apparaître un champ particulier dans la discipline de l'évaluation, souvent appelé « évaluation d'impact ». L'objectif principal de ces méthodes est d'écarter la possibilité que des facteurs autres que l'intervention à l'étude puissent expliquer un effet observé. L'effet causal d'une intervention sur un indicateur de résultat s'exprime donc comme la différence entre le résultat obtenu avec l'intervention et ce que l'on aurait pu observer sur cet indicateur sans l'intervention, ce qu'on appelle un contrefactuel. Une grande partie du raisonnement consiste donc à se demander : que se serait-il passé sans intervention ?

Toutefois, comme on le verra au cours du chapitre, répondre à cette question est souvent une tâche compliquée, et les difficultés diffèrent beaucoup selon la nature de l'intervention, sa mise en œuvre, son contexte et selon la situation d'évaluation. C'est pourquoi il n'existe pas qu'une seule méthode d'évaluation des effets, mais une diversité de méthodes à utiliser selon le diagnostic qu'une équipe d'évaluation établit pour chacune de ses missions. Ce chapitre vise à faire un tour d'horizon de cette diversité de méthodes, en commençant par celles de nature quantitative, reposant sur les statistiques, mais en présentant également des méthodes qualitatives et mixtes.

## Les méthodes quantitatives d'évaluation des effets

Les méthodes quantitatives d'évaluation des effets reposent sur la notion statistique d'inférence causale, et leur raisonnement est construit autour de la notion de contrefactuel, c'est-à-dire une estimation de la situation sans intervention. Par exemple, si un récent diplômé d'un programme de formation trouve un emploi, on se demandera si cette embauche est un résultat direct du programme, ou si l'individu aurait de toute façon trouvé du travail même en dehors de cette formation. Pour répondre à cette question, on souhaite déterminer si d'autres facteurs que la formation ont pu jouer un rôle, et dans quelle mesure la formation est l'élément déterminant de l'embauche (Hempel et Fiala, 2011).

La façon la plus usuelle d'obtenir un contrefactuel est de former ce que l'on appelle **un groupe témoin**. Un groupe témoin est un groupe qui possède les mêmes caractéristiques que le groupe exposé à l'intervention (aussi appelé groupe de traitement), excepté le fait que les sujets du groupe témoin n'ont pas bénéficié de l'intervention. Les résultats sur le groupe témoin nous renseignent donc sur la situation sans intervention, le contrefactuel. Les termes «groupe témoin» et «groupe de contrôle» sont souvent utilisés indifféremment, bien que, à proprement parler, le second terme soit applicable uniquement dans le cadre d'évaluations expérimentales, sur lesquelles la prochaine section reviendra (Hempel et Fiala, 2011). En utilisant ce contrefactuel, on peut alors faire la différence entre l'effet brut d'une intervention et son effet net : l'effet brut est l'ensemble des changements observés sur un indicateur cible, alors que l'effet net est celui causé par l'intervention et elle seule, comme illustré dans la figure ci-après. Si l'on reprend l'exemple de la formation développé plus haut, l'effet brut consisterait à mesurer combien d'individus formés ont obtenu un emploi après la formation, mais pour obtenir l'effet net, il faudrait retrancher de ce nombre le total d'individus qui auraient obtenu un emploi pendant cette période, même s'ils n'avaient pas suivi la formation.



**Figure 1**

L'effet brut et l'effet net d'une intervention

Source : Anstett, 2019 pour l'ENAP

## Le devis expérimental et ses applications

### Principes du devis expérimental

Le principe des devis expérimentaux, communément appelés essais contrôlés randomisés (ou en anglais *randomized controlled trials* – RCT) ou modèles par assignation aléatoire, est de fournir une estimation précise des effets nets des interventions, en répliquant les méthodes utilisées pour les essais cliniques (Mayneris, 2009) on compare des gens qui ont bénéficié d'un traitement – dans le cadre d'un essai clinique, ce sera un nouveau médicament – et des gens qui n'en ont pas bénéficié.

Pour que cette procédure fonctionne, les deux groupes (de traitement et témoin, ou dans ce cas expérimental et contrôle) doivent être équivalents en tout point, si ce n'est de leur participation à l'intervention. On peut donc dire qu'ils ont les mêmes caractéristiques sur trois plans (Gertler *et al.*, 2011) :

1. Être identiques en termes de caractéristiques observables et non observables. Les caractéristiques observables font référence à l'âge, au sexe, au niveau d'éducation, au statut socio-économique, aux caractéristiques familiales, à la situation au regard de l'emploi, etc. Les caractéristiques non observables comprennent la motivation, l'intérêt, les préférences, le degré de soutien familial et d'autres facteurs. Si les sujets du groupe de traitement ne doivent pas tous être identiques aux sujets du groupe témoin, les deux groupes devraient être en moyenne similaires.
2. Réagir au programme de la même façon. Par exemple, la probabilité que les résultats, tels que le développement des compétences ou l'amélioration du revenu, devrait être la même pour les sujets du groupe de traitement que pour ceux du groupe témoin.
3. Bénéficier de manière égale d'autres interventions. Par exemple, les deux groupes devraient avoir un accès égal aux autres services d'aide fournis par les autorités locales, les ONG, etc.

Lorsque les conditions énumérées ci-dessus sont assurées, seule l'existence de l'intervention expliquera toute différence dans les résultats. Dans ce cas, l'effet de causalité du programme peut être prouvé et on parlera alors de l'attribution : « la confirmation d'une relation causale entre les changements observés (ou que l'on s'attend à observer) et une action spécifique » (OCDE, 2002). Si, en revanche, le groupe témoin diffère de manière significative du groupe de traitement, nous sommes face à un biais de sélection (ou erreur systématique), qui invalidera nos mesures de l'impact. Ces biais de sélection renvoient au fait que des différences profondes entre le groupe de traitement et le groupe témoin expliquent la raison pour laquelle nous observons des résultats différents. Les biais de sélection apparaissent souvent lorsque le groupe témoin est composé de personnes qui ne remplissent pas les conditions requises pour bénéficier du programme (conditions fondées sur des caractéristiques observables) ou qui ont choisi de ne pas y participer (pour des raisons non observables) (Hempel et Fiala, 2011).

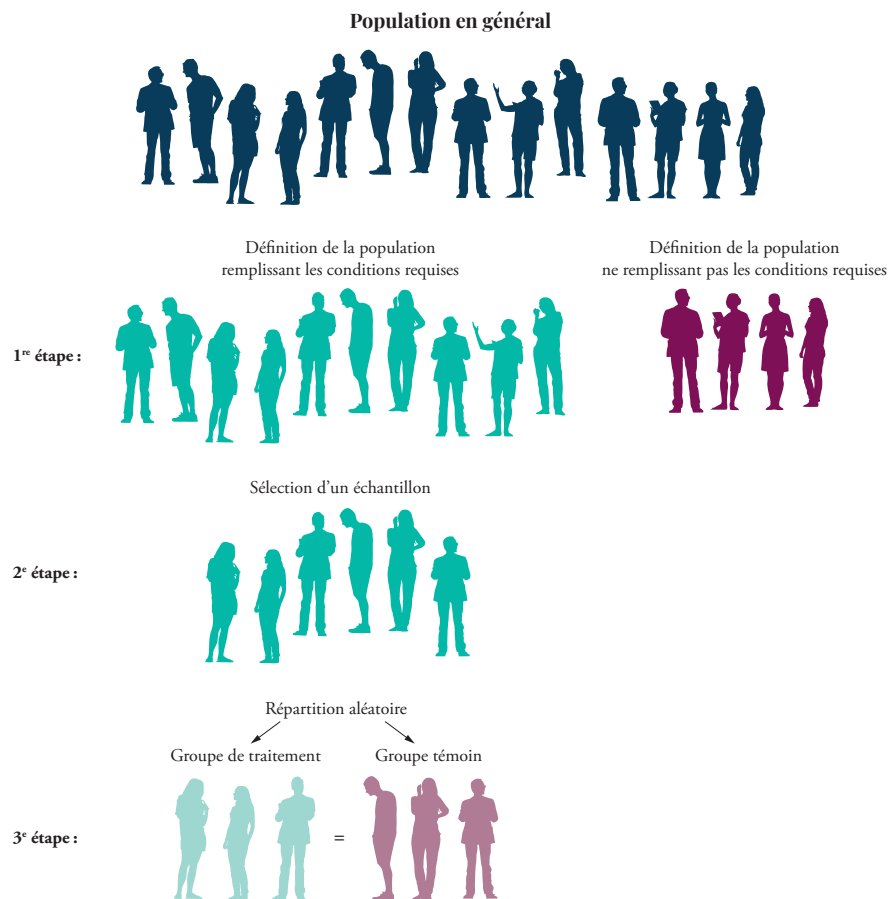
Le moyen le plus sûr d'éviter les biais de sélection, et qui est utilisé dans le devis expérimental, est une assignation aléatoire dans les groupes de traitement et témoin avant la mise en œuvre de l'intervention. Lorsque les deux groupes sont constitués au hasard et en assez grand nombre, dans la même population admissible, ils auront alors des caractéristiques moyennes similaires (à l'exception qu'un

groupe a été soumis à l'intervention et l'autre, non). Dans les procédures expérimentales telles que les essais cliniques, l'assignation aléatoire est donc devenue « l'étalon-or » des modèles d'évaluation des effets, car elle permet de résoudre le défi de l'attribution avec un niveau élevé de certitude.

### Réalisation pratique du devis expérimental

L'assignation aléatoire des participants dans le cadre d'un devis expérimental peut être réalisée en trois étapes (Hempel et Fiala, 2011) :

- 1<sup>re</sup> étape : définition de la population remplissant les conditions requises pour l'intervention (la population admissible);
- 2<sup>e</sup> étape : sélection d'un échantillon d'évaluation représentatif de la population (on utilise une méthode probabiliste). Ce sont les unités sur lesquelles on collectera des données;
- 3<sup>e</sup> étape : assignation aléatoire dans les deux groupes (groupe de traitement et groupe témoin).



**Figure 2**

Les étapes de réalisation du devis expérimental

Source : Adapté de Hempel et Fiala, 2011, p. 86

Une fois la sélection des groupes effectuée, il est possible de réaliser deux modèles de mesures selon que l'on souhaite (ou que l'on puisse) disposer de données préintervention (type M1) ou non (type M2). Ces données ne sont pas essentielles, mais pourraient permettre une vérification de l'équivalence entre les deux groupes avant l'intervention sur un certain nombre de caractéristiques observables.

Les devis sont souvent représentés par des tableaux qui reprennent :

- En lignes, les groupes sur lesquels des mesures sont effectuées (dans notre cas le groupe de traitement et le groupe témoin).
- En colonnes, la temporalité des mesures et de l'intervention avec la légende suivante :
- **X** : exposition à l'intervention
- **E** (ligne du haut) : mesures pour le groupe de traitement ou groupe expérimental
- **C** (ligne du bas) : mesures pour le groupe témoin ou groupe de contrôle
- **R** : indique que l'on a réalisé une sélection aléatoire ou randomisée des groupes

**Le modèle classique : mesure avant et après (type M1)**

	Mesure avant	Exposition à l'intervention	Mesure après
<i>Groupe de traitement (R)</i>	E <sub>1</sub>	X	E <sub>2</sub>
<i>Groupe témoin (R)</i>	C <sub>1</sub>		C <sub>2</sub>

$$\text{Calcul de l'effet net} = (E_2 - E_1) - (C_2 - C_1)$$

Note : E<sub>1</sub> et C<sub>1</sub> sont, en moyenne, identiques en tous points avant l'intervention X. Ainsi, le calcul de l'effet net peut être simplifié en ne faisant que E<sub>2</sub> - C<sub>2</sub>. C'est pourquoi on utilise aussi beaucoup le modèle M2.

**Le modèle simplifié : mesure après (type M2)**

	Mesure avant	Exposition à l'intervention	Mesure après
<i>Groupe expérimental (R)</i>		X	E <sub>2</sub>
<i>Groupe témoin (R)</i>			C <sub>2</sub>

$$\text{Calcul de l'effet net} = E_2 - C_2$$



### Les avantages et les limites du devis expérimental

Les procédures expérimentales ont été pensées pour des conditions de laboratoire, ce qui malheureusement ne s'applique pas beaucoup à la réalité d'implantation des interventions publiques. Ainsi, le devis expérimental n'est pas toujours applicable dans les missions d'évaluation malgré ses avantages. On considère que l'utilisation d'un devis expérimental est notamment possible dans l'évaluation des interventions si (Gertler *et al.*, 2011) :

- a. L'évaluation est prévue avant la mise en œuvre de l'intervention, ce qui laisse donc la possibilité d'effectuer une sélection aléatoire de la population qui y sera exposée. Les évaluations finales ou *ex post* arrivent souvent trop tard pour que l'on puisse reconstituer les conditions du devis expérimental.
- b. Le nombre d'unités admissibles est supérieur au nombre de places disponibles dans l'intervention. Si, pour des raisons de budget, la demande pour l'intervention dépasse l'offre, alors le fait de mettre de côté des non-bénéficiaires pour des raisons évaluatives est moins un problème. On peut alors soutenir que de sélectionner des bénéficiaires au hasard peut même répondre à une demande d'équité.
- c. L'intervention doit être progressivement étendue pour couvrir l'ensemble de la population admissible.

Lorsque ces conditions sont réunies, il est alors possible de constituer un groupe témoin. Mais les interventions publiques, en dehors de projets pilotes, ne sont que rarement conçues de cette façon, et on applique généralement d'autres critères de sélection pour sélectionner les bénéficiaires, en préférant choisir la population qui a le plus besoin de services, plutôt que de sélectionner des personnes au hasard. Il est donc difficile de dissocier les méthodes expérimentales d'un questionnement éthique : la possibilité de réaliser ce type d'évaluation est-elle une raison suffisante pour accepter de laisser de côté des non-bénéficiaires, choisis au hasard ?

Toutefois, lorsque ces conditions sont présentes, la méthode du devis expérimental offre plusieurs avantages considérables, dont (Hempel et Fiala, 2011) :

- Elle permet l'estimation la plus précise du contrefactuel, et par conséquent la mesure la plus rigoureuse statistiquement de l'effet net d'une intervention, tout en permettant de mesurer la magnitude des effets;
- C'est de loin la méthode d'évaluation la plus simple sur le plan analytique : l'effet de l'intervention est simplement la différence moyenne entre les résultats obtenus dans le groupe de traitement et ceux obtenus dans le groupe témoin;
- Puisqu'elle doit être prévue dès le lancement de l'intervention, elle peut être conçue non seulement pour mesurer l'impact moyen d'une intervention, mais aussi pour comparer l'efficacité de différents éléments et de différentes durées d'intervention, sur une diversité de groupes de populations, etc.

Cependant, il est souvent difficile de concevoir et de mettre en œuvre un devis expérimental à cause des problèmes suivants :

- Des questions éthiques concernant l'exclusion au hasard de certaines personnes de l'intervention, comme expliqué plus haut. D'autres méthodes viennent toutefois répliquer les conditions du devis expérimental pour des méthodes de sélection de la population bénéficiaire différentes (double différence, discontinuité de la régression par exemple);
- La portée du programme peut signifier qu'il n'y a pas de groupes de populations non exposées en tant que telles, avec une intervention de large portée – par exemple un prêt d'ajustement ou des interventions administrées à un niveau national, comme une campagne de sensibilisation ou une politique sectorielle (Hempel et Fiala, 2011). Certaines méthodes, telles que les séries temporelles, peuvent toutefois être utilisées en l'absence de groupe témoin;
- L'exigence de nombreuses données individuelles : outre le coût très élevé de collecte que ces devis peuvent représenter, il y a souvent une non-faisabilité logistique – en particulier dans les pays en développement – car il faut suivre les individus sur une longue période, en particulier des non-bénéficiaires. On parle alors de biais d'attrition ou de mortalité expérimentale;
- L'implication nécessaire des évaluateurs et des évaluatrices au démarrage des interventions, alors que l'habitude est beaucoup plus souvent prise de déclencher des évaluations à mi-parcours ou des évaluations finales.

## Devis quasi expérimentaux

Les difficultés de réalisation du devis expérimental ont conduit les évaluateurs et évaluatrices à faire appel à des méthodes alternatives. L'estimation du contrefactuel peut notamment être possible en recréant un groupe témoin ressemblant suffisamment au groupe de traitement choisi sans assignation aléatoire, mais permettant tout de même de faire des inférences valides. On entend donc par « devis quasi expérimentaux », des méthodes qui s'appliquent dans des situations où toutes les conditions du devis expérimental ne sont pas réunies, mais où l'on souhaite tout de même répliquer sa logique de calcul (différence mesurée sur une variable pour un groupe exposé à une intervention et un groupe non exposé).

### La double différence

La méthode de la double différence (DD), ou de la divergence dans la différence, ou encore des différences de différences (en anglais *differences in differences* – DID) est très couramment utilisée. Elle vise à comparer les différences de résultats au fil du temps entre une population exposée à un programme (le groupe de traitement, comme celui utilisé dans le devis expérimental) et les résultats d'un autre groupe, recréé *a posteriori* n'y étant pas exposé. Les deux groupes ne sont donc pas issus de la même population admissible avant l'intervention, on souhaite plutôt trouver un groupe témoin acceptable parmi un ensemble de non-bénéficiaires.

Cette méthode requiert de disposer de données de référence collectées avant le début de l'intervention. Pour le calcul, on effectue une première soustraction « après – avant » pour le groupe de traitement. C'est la première différence, qui correspond à l'effet brut de l'intervention. Pour écarter les effets liés à des facteurs externes, on effectue une deuxième différence, cette fois-ci pour un groupe qui n'a pas été exposé à l'intervention, mais qui est sujet aux mêmes facteurs externes, ce qui permettra ensuite de dégager l'effet net de l'intervention (Gertler, *et al.*, 2011).

	Mesure avant	Exposition à l'intervention	Mesure après	2 <sup>e</sup> différence – comparaisons avant et après intervention
<i>Groupe de traitement</i>	E 1	X	E 2	E2-E1
<i>Groupe témoin</i>	C 1		C 2	C2-C1
1 <sup>re</sup> différence – Comparaisons avec et sans intervention	C1-E1		C2-E2	Double différence = (C2-E2) – (C1-E1) = (E2-E1) – (C2-C1)

$$\text{Calcul de l'effet net} = (E_2 - E_1) - (C_2 - C_1)$$

Généralement moins solide que les devis expérimentaux qu'elle essaie de répliquer, la méthode DD ne requiert pas de choisir les règles d'assignation de l'intervention en fonction de l'évaluation. Les gestionnaires de l'intervention peuvent donc choisir des bénéficiaires selon d'autres critères. Sans la condition de l'assignation aléatoire, l'application de cette méthode est donc facilitée, à condition qu'il existe un bassin de population non exposée. En revanche, même si les tendances sont équivalentes entre les deux groupes avant l'intervention, l'estimation de l'effet net peut présenter un biais. En effet, la méthode DD fait le lien entre l'intervention et toutes les différences que l'on va mesurer entre le groupe de traitement et le groupe de comparaison. S'il existe d'autres facteurs influençant la différence de tendances entre les deux groupes, l'estimation sera non valide ou biaisée (Gertler, *et al.*, 2011). Puisqu'il s'agit d'un contrefactuel choisi par l'équipe d'évaluation, il lui revient donc d'argumenter et de justifier ce choix au lecteur.

### La méthode d'appariement

L'appariement « consiste à utiliser de grandes bases de données et des techniques statistiques complexes pour générer le meilleur groupe de comparaison artificiel possible pour un groupe de traitement donné » (Gertler *et al.*, 2011, p. 104). La méthode consiste à s'assurer que, sur un ensemble de variables déterminées (dont la variable d'effet), le groupe de comparaison est similaire (idéalement identique) au groupe de traitement avant l'application de l'intervention. Comme pour la méthode de la double différence, on souhaite répliquer les conditions du devis expérimental, la différence ici est que le groupe témoin n'est pas simplement choisi par l'évaluateur ou l'évaluatrice, on applique une procédure précise de sélection du groupe témoin pour augmenter la validité du devis.

	Mesure avant	Exposition à l'intervention	Mesure après
<i>Groupe de traitement (A)</i>	E <sub>1</sub>	X	E <sub>2</sub>
<i>Groupe témoin (A)</i>	C <sub>1</sub>		C <sub>2</sub>

**A :** indique que l'on a réalisé une sélection appariée pour constituer le groupe témoin.

$$\text{Calcul de l'effet net} = (E_2 - E_1) - (C_2 - C_1)$$

On distingue deux méthodes d'appariement : l'appariement par paires et l'appariement statistique (méthode des scores de propension).

L'appariement par paires se réfère à la sélection du groupe de comparaison en fonction de critères spécifiques de similitude. Dans l'appariement par paires on essaie, idéalement, de trouver des personnes ou des familles similaires, une à la fois. Cependant, en général, on se contente d'essayer de parvenir à des groupes comparables qui sont similaires en moyenne comme on le fait en procédant par randomisation.

L'appariement statistique ou la méthode des scores de propension (en anglais *propensity score matching*–PSM) cherche à identifier un groupe d'individus avec les mêmes caractéristiques observables que les participants à l'intervention. Pour ce faire, on calcule, à l'aide d'un modèle de régression, des scores de propension (probabilités de participer à l'intervention, fondées sur les similarités entre participants et non-participants) qui servent à appairer participants et non-participants.

Les méthodes d'appariement présentent une alternative lorsque le devis expérimental ne peut pas être utilisé, car elle ne nécessite pas que l'évaluateur ou l'évaluatrice ait le contrôle direct sur la sélection des individus dans l'intervention. Elles peuvent être appliquées rétrospectivement, si les données sont disponibles, et peuvent être utilisées pour évaluer des effets multiples, ce qui constitue des avantages logistiques importants.

Cependant, ce sont des procédures qui requièrent des bases de données intégrées et unifiées sur les individus exposés à l'intervention, mais également sur un bassin plus large de non-participants, ce qui est dans la réalité, plus rare. Par ailleurs, on ne peut faire un appariement que sur des caractéristiques observables. Or, de nombreuses caractéristiques non observables, telles que la motivation, peuvent être la cause d'une différence entre groupe de traitement et groupe témoin (Gertler *et al.*, 2011).

### Modèles à série temporelle ou modèles chronologiques

Les modèles à série temporelle ou modèles chronologiques sont caractérisés par une série de mesures échelonnées dans le temps, à la fois avant et après l'exposition au programme. Les modèles chronologiques pour lesquels il n'existe que quelques mesures avant et après le programme sont vulnérables aux mêmes obstacles influant sur la validité interne que les modèles à mesures uniques correspondants. À l'inverse, une série complète de mesures avant et après le programme permet à l'évaluateur ou l'évaluatrice d'éliminer un grand nombre de ces obstacles, en analysant les tendances antérieures et postérieures au programme (Secrétariat du Conseil du Trésor du Canada, 1998).

On peut distinguer deux modèles chronologiques : le modèle de base à série temporelle et le modèle à série temporelle avec groupe témoin non équivalent.

#### Modèle de base à série temporelle

Les effets d'un programme donné sont déterminés en fonction du changement qui se manifeste dans la série de mesures prises avant et après l'exposition à l'intervention. On réalise une projection (un calcul statistique qui permet de prédire le futur à partir de données sur le passé) pour connaître le contrefactuel. Plutôt que de se servir d'un contrefactuel mesuré auprès de non-bénéficiaires, on se sert d'un contrefactuel calculé.

	Mesure avant	Exposition à l'intervention	Mesure après
<i>Groupe de traitement</i>	...E <sub>1</sub> , E <sub>2</sub> , E <sub>3</sub> , E <sub>4</sub>	X	E <sub>5</sub> , E <sub>6</sub> , E <sub>7</sub> , E <sub>8</sub> ...

**Modèle à série temporelle avec groupe témoin non équivalent**

	Mesure avant	Exposition à l'intervention	Mesure après
<i>Groupe de traitement</i>	...E <sub>1</sub> , E <sub>2</sub> , E <sub>3</sub> , E <sub>4</sub>	X	E <sub>5</sub> , E <sub>6</sub> , E <sub>7</sub> , E <sub>8</sub> ...
<i>Groupe témoin</i>	...C <sub>1</sub> , C <sub>2</sub> , C <sub>3</sub> , C <sub>4</sub>		C <sub>5</sub> , C <sub>6</sub> , C <sub>7</sub> , C <sub>8</sub> ...

Ici, on choisit le groupe témoin en souhaitant qu'il soit soumis aux mêmes facteurs externes que le groupe de traitement afin qu'un changement observé soit attribuable à l'intervention. Comme le groupe témoin est non équivalent, il faut toutefois que les groupes se ressemblent suffisamment en ce qui concerne les caractéristiques étudiées. Si c'est le cas, un modèle chronologique peut ainsi se révéler très intéressant et il permet d'accroître la validité d'un devis à double différence (Secrétariat du Conseil du Trésor du Canada, 1998).

Un avantage majeur des modèles chronologiques fondés sur des données de série temporelle fiables consiste en l'élimination de nombreux obstacles à la validité interne. Divers effets du programme dépendant du facteur temps peuvent être analysés grâce à ce modèle, d'autant plus que leur caractère longitudinal permet également d'étudier la temporalité des effets ou encore leur durabilité dans le temps (Secrétariat du Conseil du Trésor du Canada, 1998).

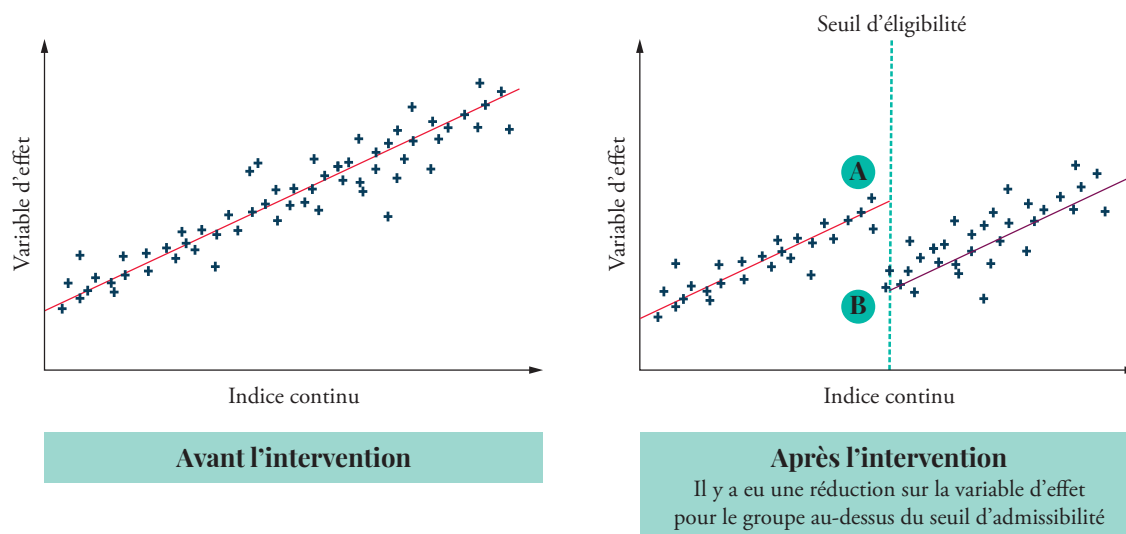
Une des difficultés importantes est que l'on ne dispose pas toujours de données fiables pour mener l'analyse à série temporelle qui s'impose, surtout dans le contexte de développement. Ces modèles posent de nombreux problèmes de données, et dans la réalité les séries de données disponibles ne sont pas toujours assez longues et stables dans le temps (même outil de mesure ou questions de sondage posées).

**Le modèle de discontinuité de la régression**

Cette quatrième méthode quasi expérimentale convient aux interventions pour lesquelles un indice d'admissibilité continu est établi (comme un niveau de revenu, les résultats à un examen ou l'âge) et un seuil est clairement défini pour distinguer les bénéficiaires des non-bénéficiaires (Baker, 2000).

Le postulat dans les procédures d'évaluation par discontinuité de la régression (ou indice d'admissibilité) est que les personnes qui obtiennent un score juste au-dessus ou juste en dessous d'un seuil préalablement fixé ne sont pas très différentes les unes des autres, ou du moins la différence peut être continue à travers les scores (Hempel et Fiala, 2011). On se base donc sur le fait qu'avec une intervention publique, on crée une discontinuité (autour du seuil d'admissibilité), là où il avait une continuité, la cassure étant une bonne estimation de l'effet de l'intervention.

Les participants dont le score est juste au-dessus du seuil d'admissibilité sont en effet considérés comme très similaires à ceux dont le score est juste en-dessous du seuil d'admissibilité. Le choix d'un point particulier sur l'indice continu crée une coupure, ou une discontinuité, dans l'admissibilité au programme. « Étant donné que les participants qui se situent juste au-dessus du seuil sont très similaires à ceux qui sont juste en dessous, à la différence près qu'ils ne bénéficient pas de l'intervention, ils peuvent être utilisés comme groupe de comparaison. Autrement dit, les participants non éligibles à l'intervention, mais proches du seuil d'éligibilité seront utilisés comme groupe de comparaison pour estimer le contrefactuel » (Gertler, *et al.*, 2011).



Calcul de l'effet net = A-B

**Figure 3**

La discontinuité de la régression

Source : Anstett, 2018 pour l'ENAP

Le modèle de discontinuité de la régression utilise les règles de ciblage existantes et n'impose pas un changement dans la mise en œuvre de l'intervention. Il présente l'avantage de mesurer des estimations sans erreur systématique pour les participants ayant des notes proches du seuil. L'absence de la sélection aléatoire, et donc l'exclusion d'une partie de la population, rend le modèle politiquement plus acceptable, et logistiquement plus simple dans le cas spécifique des interventions avec seuil d'admissibilité. Cependant, le modèle ne permet la mesure d'impact de l'intervention que localement aux alentours du seuil d'admissibilité, et on ne peut pas calculer l'effet moyen du traitement pour tous les participants. Par ailleurs, pour que la méthode fonctionne, les règles d'admissibilité doivent être appliquées strictement dès le démarrage de l'intervention (il ne faut pas qu'il y ait de contagion entre les deux groupes). Or, on remarque qu'en réalité, des négociations ou des adaptations des participants potentiels peuvent venir fausser un peu les critères de sélection, ce qui entache ensuite la validité des résultats (Hempel et Fiala, 2011).

Il existe encore une quantité d'autres méthodes quasi expérimentales, plus ou moins dérivées de la méthode du devis expérimental, pour faire face à toutes les situations d'interventions publiques, et de disponibilités de données. Des références supplémentaires sont proposées aux lecteurs dans la section « Pour aller plus loin », et de nombreuses publications scientifiques et pratiques sont disponibles en ligne pour se rendre compte de la diversité des méthodes quantitatives appliquées en évaluation.

### Encadré 1 – Évaluation d'impact du « Programme d'amélioration de la productivité de l'agriculture irriguée du Sindh » (Source : Mission d'évaluation réalisée par A. Ben Abdallah, 2017).

Le fleuve Indus, ou le Sindh, représente la source principale de l'irrigation de la région exposée au programme. Jusqu'en décembre 2013, sur 46 699 cours d'eau dans le Sindh, seulement 29 479 cours d'eau avaient été aménagés dans le cadre de divers programmes financés localement et par des organismes donateurs. Il restait donc environ 17 220 cours d'eau à aménager, car le revêtement des cours d'eau offre des avantages importants, dont l'augmentation marquée de l'efficacité de transport de l'eau. Ainsi, l'eau économisée s'ajoute à l'augmentation des potentiels d'irrigation, ce qui a pour effet souhaité une production agricole améliorée dans un contexte de raréfaction des sources d'eau.

Le programme évalué est une collaboration entre le gouvernement du Sindh et la Banque mondiale, et vise le revêtement de 5 500 cours d'eau sur une période de six ans. La démarche préconisée cherche également à impliquer les agriculteurs pour garantir la pérennité des installations. Une coopérative de six propriétaires pour chaque cours d'eau doit, par exemple, être formée au préalable. Pour chaque cours d'eau, une procédure d'appui en plusieurs étapes est établie afin de garantir l'adhésion et la participation de ces coopératives aux travaux et à l'exploitation équitable de la source d'eau.

La réalisation des travaux est répartie sur les six années comme suit :

Année	2015	2016	2017	2018	2019	2020	Total
Cibles d'aménagement de cours d'eau	300	1100	1100	1000	1000	1000	5500

Le programme inclut également d'autres activités connexes telles que l'installation de nouveaux systèmes d'irrigation, la construction d'abris contre les inondations, l'accompagnement vers des pratiques agricoles modernes (utilisation d'engrais et de pesticides, défonçage profond, expansion du nivellement de précision) ou la création de guichets d'information et de formation. Le nombre total des bénéficiaires directs et indirects du programme sont estimés à plus de deux millions d'individus.

#### Méthode choisie

Une étude de base (*Baseline*) est réalisée avant le lancement des activités du programme. Un système informatisé de suivi-évaluation est également mis en place. Bien que le nombre de cours d'eau admissibles au programme dépasse la cible du programme, le choix d'un devis expérimental a été écarté pour plusieurs raisons, notamment en raison de l'incapacité d'identifier d'une manière exacte l'ensemble de la population d'agriculteurs aidés par les aménagements de chacun des cours d'eau.

Afin de constituer le contrefactuel, l'équipe de l'évaluation a donc proposé un modèle quasi expérimental combinant deux modèles : la méthode des doubles différences combinée à l'approche séquentielle (*pipeline*), qui est une autre variante des devis quasi expérimentaux. Il s'agit de sélectionner un échantillon aléatoire de ménages d'agriculteurs parmi les bénéficiaires de l'année *N* qui servirait de groupe de traitement. Un autre échantillon de ménages présentant



des caractéristiques similaires est constitué au hasard parmi ceux qui peuvent bénéficier de l'intervention l'année suivante  $N+1$  pour constituer le groupe témoin. La taille de l'échantillon a été estimée à 10 % de l'ensemble des participants au programme. Par ailleurs, on utilise une répartition des bénéficiaires en trois niveaux : début, milieu et fin de cours d'eau, et cette classification est prise en considération lors de la sélection afin d'améliorer la représentativité de l'échantillon. Ce choix de méthode a l'avantage de permettre une meilleure équivalence entre les deux groupes. Cependant, il exige l'adaptation de plan de mise en œuvre afin de permettre d'anticiper l'identification des participants au moins sur deux années successives. De sérieux risques de biais et de validités sont à considérer.

## Devis implicites

Toutes les évaluations d'effet ne sont cependant pas toujours complexes, et dans certains cas, l'attribution des effets à une intervention est évidente, et il n'y a pas lieu de mettre en place des méthodes complexes, puisque les menaces à la validité et les facteurs externes sont négligeables. Prenons par exemple le cas d'une formation de courte durée, dont l'effet principal immédiat est l'augmentation des connaissances des participants. La différence entre les résultats d'un test de connaissances effectué avant et après la formation pourrait tout à fait suffire à comprendre le changement apporté par la formation, après une vérification que d'autres facteurs ne viennent pas influencer sur cette différence. Dans ce cas, il serait même certainement étrange de vouloir constituer un groupe de contrôle, car sans formation et sur un court laps de temps, il y a peu de chances que les connaissances d'un groupe de contrôle augmentent significativement. L'utilisation d'un contrefactuel n'est donc pas nécessaire, car la variation de l'indicateur d'effet serait nulle. C'est pourquoi on parle, dans ce cas, de l'utilisation de devis « implicites » ou « pré-expérimentaux », parce que le groupe de contrôle n'est volontairement pas utilisé, car on peut plutôt soutenir logiquement qu'un effet mesuré est attribuable à une intervention. L'évaluateur ou l'évaluatrice justifie ainsi son choix et doit convaincre son lecteur de la validité de son raisonnement, plutôt que d'appuyer la validité de sa méthode sur le respect d'une procédure précise comme pour le devis expérimental.

Plusieurs simplifications et configurations sont possibles, avec ou sans groupe de contrôle.

### Devis implicites sans groupe de contrôle

Les devis implicites sans groupe de contrôle sont les plus courants, on fait donc une ou deux mesures auprès d'un groupe expérimental uniquement, choisi au sein de la population exposée à l'intervention, ce qui peut se représenter de la façon suivante :

	Ex ante	Intervention	Ex post
<i>Groupe expérimental</i>	<b>E 1</b>	<b>X</b>	<b>E 2</b>
<i>Groupe contrôle</i>			

$$\text{Calcul de l'effet net} = E_2 - E_1$$



	Ex ante	Intervention	Ex post
<i>Groupe expérimental</i>		X	E <sub>1</sub>
<i>Groupe contrôle</i>			

### Calcul de l'effet net = E<sub>1</sub>

Dans les deux cas, puisque l'on n'utilise pas de groupes de contrôle, on fait l'hypothèse qu'il n'y a pas de facteurs externes qui jouent sur l'indicateur mesuré, et que la quantité d'effet calculée est logiquement attribuable à l'intervention.

Si l'on ne fait qu'une seule mesure après, c'est que l'on pose l'hypothèse que la situation de départ est égale à zéro, et que le niveau mesuré sur l'indicateur d'effet est entièrement causé par l'intervention. C'est le cas lorsque l'on met en place une activité nouvelle, par exemple si un programme de tri des déchets est mis en place dans une ville alors qu'il n'y en avait pas du tout avant, et que l'on veut connaître la quantité de déchets qui est dirigée vers le tri grâce au programme, il est très probable qu'une seule mesure soit nécessaire.

Si l'on fait deux mesures, une avant et une après l'intervention, alors on considère que l'indicateur d'effet ne part pas de zéro, mais que la différence entre les deux mesures n'est attribuable qu'à l'intervention évaluée. C'est le cas du test de connaissances abordé un peu plus haut.

### Devis implicites avec groupe de contrôle

Certains devis implicites font tout de même appel à la notion de contrefactuel. La plupart du temps, il n'y aura pas de sélection aléatoire ou par appariement, mais il s'agira d'un contrefactuel choisi, comme dans la méthode de la double différence abordée à la section précédente.

Toutefois dans ce cas, en plus de faire l'hypothèse que les groupes sont équivalents, on considère également que les deux groupes partent du même point de départ, ce qui permet de simplifier le calcul de l'effet net, en éliminant la nécessité de faire une mesure avant, comme dans le cadre du devis expérimental. L'effet net est donc calculé de la même façon, soit :  $\text{Effnet} = E_1 - C_1$

	Ex ante	Intervention	Ex post
<i>Groupe expérimental</i>		X	E <sub>1</sub>
<i>Groupe contrôle</i>			C <sub>1</sub>

### Calcul de l'effet net = E<sub>1</sub> - C<sub>1</sub>

### Choix d'une méthode quantitative

Cette diversité de méthodes quantitatives peut paraître accablante pour les évaluatrices ou évaluateurs débutants, mais quelques questions peuvent guider la pose d'un diagnostic sur une situation à évaluer et, ultimement, peuvent aider à choisir la bonne méthode d'évaluation.

Premièrement, les méthodes quantitatives reposent toutes sur la prémisse qu'un indicateur quantitatif (ou un petit nombre d'indicateurs quantitatifs) représente bien l'effet que l'on cherche à évaluer, et que la seule connaissance de la quantité d'effets attribuables à une intervention suffit à répondre au besoin d'information à la base de l'évaluation. La section suivante sur les méthodes qualitatives et mixtes d'évaluation des effets reviendra sur les situations où cette prémisse n'est pas présente.

Lorsque cette prémisse est présente, on peut ensuite se poser un certain nombre de questions pour choisir le devis qui s'appliquera le mieux à la mission d'évaluation :

- Quelle est la taille de l'offre de l'intervention par rapport à la taille de la demande pour cette intervention? Avons-nous potentiellement des bénéficiaires et des non-bénéficiaires? C'est une condition nécessaire à la mise en place d'un groupe témoin, qui est une composante de nombreux devis. Les devis chronologiques restent toutefois une option.
- Peut-on effectuer une assignation des bénéficiaires avant l'intervention? Peut-on la faire uniquement pour des raisons évaluatives (sélection aléatoire pour un devis expérimental), ou d'autres considérations doivent-elle prévaloir? En l'absence d'assignation aléatoire, on peut utiliser les méthodes quasi expérimentales.
- Si une autre sélection est faite, peut-on tout de même reconstituer un bon groupe de contrôle à partir des non-bénéficiaires? Sont-ils notamment soumis aux mêmes conditions externes, et ont-ils les mêmes caractéristiques que les bénéficiaires?
- De quel historique de données dispose-t-on sur l'indicateur d'effet (notamment, avons-nous des mesures avant l'intervention s'il s'agit d'une évaluation de mi-parcours ou finale)? Sinon, dispose-t-on de données régulièrement collectées pour la mise en place d'une méthode chronologique? Ou d'une base de données des participants pour un appariement?
- Y a-t-il beaucoup de facteurs externes qui influent sur l'indicateur d'effet? S'ils sont négligeables, on peut penser à simplifier l'évaluation avec un devis implicite.

## Les méthodes qualitatives ou mixtes d'évaluation des effets

De nombreux auteurs ont relevé des limites importantes à propos des méthodes quantitatives d'évaluation des effets, particulièrement lorsqu'il s'agit de les appliquer à des interventions en contexte de développement. En effet, leurs théories du changement ont tendance à être complexes : elles s'attaquent généralement à des problèmes multifactoriels où de nombreux acteurs interviennent de façon plus ou moins cohérente sur les mêmes problématiques, dans un contexte évoluant rapidement et de manière peu prévisible (Bamberger, 2012; IE4D Group, 2011). Il devient alors difficile, voire impossible, d'affirmer avec certitude qu'une quantité précise d'effets est directement attribuable à une seule intervention évaluée, et ce pour diverses raisons :

- De l'incertitude sur le type d'effets à évaluer, par exemple dans le cas d'interventions innovantes ou mises en œuvre dans un contexte d'urgence;
- Des difficultés théoriques, techniques ou logistiques rendant les devis quantitatifs non applicables (indisponibilité des données ou incapacité à les collecter dans un délai raisonnable, absence d'indicateurs quantitatifs d'effets pertinents, trop d'interventions complémentaires à démêler, etc.);

- La non-pertinence de la mesure quantitative des effets : le besoin d'information des décideurs peut, par exemple, se situer plus au niveau de la compréhension fine des mécanismes permettant des changements sur une population cible grâce à une intervention (on se demande « comment? » et « pourquoi? » plutôt que « combien? » en rentrant dans la fameuse « boîte noire de l'intervention »).

Dans ce contexte, diverses méthodes d'évaluation des effets ou même de réelles approches, selon la définition citée dans le chapitre « Le choix des approches évaluatives », plus qualitatives ou mixtes, ont émergé plus récemment, s'intéressant non pas à l'**attribution** précise d'une quantité d'effets à une intervention, mais plutôt à la compréhension plus riche de la **contribution** d'une intervention à divers changements. On les regroupe généralement sous les termes de méthodes (ou approches) basées sur la théorie (d'intervention, car on documente et on remet en question les liens d'une théorie de changement), ou de méthodes liées à l'analyse de la contribution telles que présentées dans le chapitre 7 : *Le choix des approches évaluatives*. Ces méthodes se prêtent particulièrement aux interventions en contexte de développement, et plusieurs d'entre elles sont d'ailleurs mises en avant par des équipes travaillant dans ce secteur. Cette section vise à donner des exemples de méthodes qualitatives et mixtes d'évaluation des effets parmi les plus utilisées, sans pouvoir toutes les citer, puisque les différentes communautés de pratiques font émerger de nombreuses variantes et adaptations.

### Méthodes pouvant être à la base d'un système de suivi-évaluation

Comme on a pu le voir en passant en revue des méthodes quantitatives, l'évaluation des effets est souvent plus facile (disponibilité des données), mais également plus pertinente (interprétation et utilisation des résultats) lorsqu'elle fait partie d'un système plus large de suivi-évaluation (IE4D Group, 2011). Dans ce contexte, on peut citer deux méthodes d'évaluation qualitative des effets qui s'intègrent au cycle de vie d'une intervention de manière synergique. Ces deux méthodes sont donc particulièrement indiquées pour gérer des interventions complexes en gardant une préoccupation pour les résultats, même s'ils sont plus diffus, et plus difficiles à prévoir ou à cerner.

#### Le changement le plus significatif

Le changement le plus significatif (CPS) est une méthode de suivi-évaluation participative, fondée sur le recueil de récits de changement émanant des bénéficiaires ou par les équipes travaillant quotidiennement auprès des populations ciblées, puis sur la sélection des histoires les plus « significatives », c'est-à-dire en lien avec les finalités souhaitées pour l'intervention (Davies et Dart, 2005). L'intérêt de cette méthode est le lien qu'elle permet de faire entre le terrain et les décideurs, puisque différents panels de sélection de récits de changement significatifs doivent être mis en place, et ce processus de sélection peut aider à calibrer les activités réalisées et à en faire le tri, pour qu'elles soient les plus pertinentes et efficaces (on poursuit les activités qui sont liées aux récits de changement intéressants, et on arrête ou réoriente les activités qui sont liées à des changements non ciblés). Elle a également du succès parce qu'elle se fonde sur les réalités vécues par les populations ciblées par l'intervention évaluée et que leurs points de vue sont placés au centre d'un protocole de suivi tout au long de l'intervention, alors que la voix des populations est parfois rare au milieu d'indicateurs plus quantitatifs. Le CPS est d'ailleurs souvent appelé « évaluation sans indicateurs », car ce sont les récits de changement qui sont au centre de l'analyse, ce qui est particulièrement utile pour des interventions innovantes pour lesquelles les résultats sont diffus ou peu compris. Le CPS est une méthode mixte, bien qu'à dominante qualitative, puisqu'elle peut inclure une étape de quantification des récits (analyse de

l'étendue d'un type de récits sur l'ensemble des bénéficiaires) pour des besoins d'évaluation. Il faut toutefois noter qu'il s'agit d'une méthode qui peut paraître lourde, si d'autres mécanismes de suivi sont déjà en place, mais elle est particulièrement indiquée dans des cas où on ne dispose pas d'une théorie de changement claire, comme dans l'exemple présenté dans l'encadré 2.

**Encadré 2 – Exemple de l'évaluation d'un programme humanitaire (Source : Mission d'évaluation réalisée par E. Anstett, 2020).**

La technique du changement le plus significatif a été utilisée pour renouveler la gestion d'un programme de soutien à des populations pendant une situation de crise sécuritaire et migratoire. Après quelques années de mise en œuvre dans un contexte en constant changement, le bailleur et les différents partenaires de mise en œuvre du programme ne s'entendaient plus sur une vision commune, et les activités de terrain étaient devenues très diversifiées, ce qui rendait la gouvernance d'ensemble chaotique et le suivi, impossible.

L'évaluation à mi-parcours, demandée par le bailleur pour remédier à cette situation, a été l'occasion de dresser un bilan des effets existants du programme tel qu'il était mis en œuvre par les différents partenaires, grâce au recueil de récits de changement. Des ateliers de réflexion ont par la suite été animés par l'équipe d'évaluation, en s'inspirant des procédures de sélection décrites par Davies et Dart (2005) dans le but de sélectionner les récits qui concordaient avec la mission du programme et les visions des partenaires. Le fait de discuter sur des récits réels, basés sur les mots des bénéficiaires, a désamorcé beaucoup de conflits entre les parties prenantes, qui avaient souvent été créés par des débats sémantiques, et des différences culturelles entre le bailleur et les partenaires qui n'utilisaient pas le même langage. À la fin de ce travail de sélection, le rapport d'évaluation a pu proposer une théorie de changement renouvelée et exprimée dans le vocabulaire des bénéficiaires (avec des variantes, notamment des hypothèses pour refléter les différentes zones d'implantation du programme), et des recommandations liées aux bonnes pratiques des activités qui ont contribué aux récits de changement sélectionnés. La technique du CPS a par la suite été conservée par les équipes du programme, à la demande du bailleur, comme base de leur système de suivi-évaluation, pour continuer à faire évoluer le programme.

Elle a toutefois nécessité plus de temps que prévu pour l'évaluation à mi-parcours, et d'autres questions d'évaluation qui avaient été posées à l'équipe d'évaluation ont dû être mises de côté.

### La cartographie des incidences

La cartographie des incidences est une méthode développée au Centre de recherches pour le développement international (CRDI) au Canada, et a donc été pensée pour s'adapter aux réalités du développement. Elle ne se concentre pas sur les impacts d'une intervention, mais sur une définition innovante des incidences (*outcomes*), comprises comme des « modifications du comportement, des relations, des activités ou des agissements de personnes, de groupes et d'organisations avec lesquels un programme travaille directement » (Earl *et al.*, 2002). En observant les différentes modifications de comportements de partenaires limitrophes (la population ciblée, mais aussi des organisations de la société civile, comme des associations ou des personnes d'intérêt telles que des leaders communautaires ou des responsables locaux, des parties prenantes du secteur, etc.), la méthode vise à comprendre

comment une intervention publique y contribue. Elle permet donc de voir des effets, non pas uniquement sur les bénéficiaires, mais sur l'ensemble d'un contexte. Cette méthode permet ainsi de voir les interventions de développement différemment, comme des vecteurs d'influence sur des personnes et des organisations, ce qui modifie la façon même dont on conçoit les résultats d'une intervention. C'est pourquoi cette méthode peut être utilisée plus largement que pour une évaluation des effets, car en précisant les intentions d'une intervention de manière participative selon ce nouveau point de vue (les incidences souhaitées de l'intervention), il est possible de faire un recueil régulier des réalisations et de mettre en place un réel système de suivi-évaluation. Particulièrement indiquée pour l'évaluation d'intervention visant des changements sociétaux, à long terme, dans des communautés étendues, la structure complète de la cartographie des incidences peut être un peu lourde pour des interventions qui ont des effets plus concrets. On la réserve donc aux interventions les plus complexes. De nombreux évaluateurs et évaluatrices utilisent cette méthode et cette définition des incidences, ou l'adaptent dans leur pratique, si bien qu'on peut la considérer comme une approche (au sens présenté dans le chapitre 7 : *Le choix des approches évaluatives*). Une des variantes de la cartographie des incidences (récolte des incidences) est d'ailleurs présentée dans la prochaine section comme une méthode à part entière.

### **Méthodes ne visant que l'évaluation des effets**

D'autres méthodes, à la manière des devis quantitatifs, servent plus particulièrement à répondre aux questions d'évaluation des effets et s'intègrent plus facilement à une évaluation réalisée de manière ponctuelle, souvent *a posteriori*.

### **Évaluation rapide de l'impact**

L'évaluation rapide de l'impact (ERI) est une méthode développée par Andy Rowe et qui, à l'instar des devis quantitatifs, est fondée sur la réflexion sur la situation sans intervention. Particulièrement utile lorsque le calendrier (évaluation *ex ante* ou besoin rapide d'information) ou les données ne permettent pas la mise en place d'un devis, l'ERI est « une façon structurée de recueillir des jugements d'experts sur l'impact [d'une intervention] » (Centre d'excellence en évaluation, 2019). Un panel représentatif d'experts (intervenants de l'intervention et conseillers externes) est mis à contribution pour évaluer l'impact de différents scénarios, incluant l'intervention évaluée, ce qui permet de générer des consensus. Des protocoles précis de recueils, de pondération et d'analyses des jugements d'experts sont proposés aux évaluateurs et évaluatrices pour conserver la rigueur de la méthode. Comme son nom l'indique, cette méthode est particulièrement indiquée lorsque l'on a besoin d'une évaluation rapide, notamment lorsque l'on doit éclairer une prise de décision qui doit être faite rapidement. Elle est, en ce sens, un bon outil pour les évaluateurs et évaluatrices internes, ou dans le cadre d'évaluations en continu. Elle repose toutefois sur l'existence et la disponibilité des experts, ce qui peut ne pas s'appliquer à tous les secteurs ou à toutes les situations.

### **Récolte des incidences et autres méthodes d'analyse de la contribution**

La méthode de la récolte des incidences est un processus d'analyse à rebours : on part de changements observés sur le terrain, et l'évaluateur ou l'évaluatrice cherche à comprendre si une intervention y a contribué, et de quelle façon (Wilson-Grau, 2019). Développée par Ricardo Wilson-Grau et ses collègues à la manière d'une enquête policière, elle s'appuie sur une recherche de preuves de changement ou d'incidences (dans la même définition que pour la cartographie des incidences dont elle

s'inspire). Comme la cartographie des incidences, cette méthode est particulièrement intéressante dans les contextes complexes où l'ensemble des effets possibles d'une intervention ne sont pas connus à l'avance, et permet donc de rassembler et de collecter des données (des preuves d'incidences) très diverses. On peut donc s'en servir même lorsque l'on n'a pas une théorie de changement validée. Après l'étape de collecte, l'équipe d'évaluation étudie le lien de contribution entre l'incidence et l'intervention évaluée. La différence avec la cartographie des incidences est que la récolte des incidences ne s'utilise pas comme un système de suivi-évaluation tout au long du cycle de vie d'une intervention, mais qu'elle peut être utilisée comme une méthode d'évaluation uniquement. Cette méthode, bien que récente, est de plus en plus utilisée, et une réelle communauté de pratique s'est formée, partageant de nombreux exemples d'application et de variantes. On peut ainsi citer une diversité de méthodes découlant de l'analyse de contribution, comme le protocole d'impact qualitatif QuIP (Copestake *et al.*, 2019) qui propose une procédure précise d'entrevues et de groupes de discussion. On peut également citer l'évaluation participative du développement PADev (Dietz, *et al.*, 2013), qui s'intéresse aux changements dans un secteur au complet sur une période longue, qui peut être utile pour obtenir une perspective d'ensemble sur une problématique.

### **Encadré 3 – Exemple de l'évaluation d'une politique nationale de santé publique** (Source : Mission d'évaluation réalisée par E. Anstett, 2021).<sup>1</sup>

La méthode de la récolte des incidences a été utilisée pour évaluer les effets d'un certain nombre de réformes mises en place par l'État dans le secteur de la santé, comme la gestion participative et communautaire des établissements de santé, ou les renforcements massifs des plateaux techniques.

Des exemples d'incidences qui ont pu être documentées correspondent notamment à : la participation de la société civile dans la prise de décision des établissements de santé par la formation d'associations, la réduction des déplacements nécessaires pour obtenir différents soins ou encore les signes d'amélioration de l'état de santé des populations desservies par les établissements.

La méthode de la récolte des incidences s'applique particulièrement à ce cas, car les effets de la politique n'étaient pas tous bien cartographiés et connus à l'avance. Ainsi, l'équipe d'évaluation a pu s'intéresser à diverses preuves de changement dans le secteur de la santé sans se limiter aux indicateurs suivis dans le cadre logique de la réforme. Elle est aussi intéressante du fait qu'elle a permis d'étudier le secteur sanitaire dans son ensemble, avec son important nombre d'acteurs publics étatiques, d'organisations internationales, d'organisations non gouvernementales (ONG) ou de fondations liées par exemple au secteur minier, d'acteurs privés, de groupes d'intérêts et d'associations de la société civile, de prestataires et de patients.

Enfin, un facteur facilitant la mise en place de cette méthode a été l'exhaustivité de la politique évaluée, c'est-à-dire qu'elle visait l'intégration de l'ensemble des actions publiques sur le secteur de la santé (y compris celles des partenaires de l'aide internationale). L'analyse des liens de contribution entre la politique et les incidences recueillies a ainsi été simplifiée.

1. Le rapport complet est disponible à l'adresse suivante : [http://www.bvg-mali.org/images/Rapports/individuels/NET\\_RAPP\\_EPP\\_SOCIO-SANITAIRE\\_III.pdf](http://www.bvg-mali.org/images/Rapports/individuels/NET_RAPP_EPP_SOCIO-SANITAIRE_III.pdf)



## Synthèse sur les méthodes qualitatives ou mixtes

Tout comme les méthodes quantitatives, les méthodes qualitatives ou mixtes d'évaluation des effets sont diverses, et de nombreuses équipes d'évaluation les réinventent, les hybrident quotidiennement pour trouver la solution qui fonctionne le mieux pour chaque intervention. Grâce à l'application de ces méthodes, les possibilités sont larges : les évaluateurs et évaluatrices peuvent étudier la boîte noire des interventions, et obtenir des informations qui vont plus loin que la quantité d'effets ressentis, aller étudier qui ressent des effets et pourquoi, ou encore dans quelles conditions on peut répliquer ces effets, comment améliorer une intervention pour en obtenir plus, comment une intervention contribue à un changement en complémentarité avec d'autres interventions, etc.

Toutefois, pour éviter de se perdre dans la diversité de méthodes et de possibilités, plusieurs questions peuvent guider les évaluateurs et évaluatrices dans leur diagnostic, avant de choisir la méthode qui conviendra à chacune de leurs missions :

- Le besoin d'information est-il ponctuel (souvent pour une évaluation finale), ou la méthode peut-elle informer la gestion de l'intervention sur le long terme (évaluation en continu, ou mise en place d'un système d'évaluation)? On choisira alors une méthode pouvant s'intégrer à un système de suivi-évaluation plus large si nécessaire.
- La structure de mise en œuvre du programme est-elle simple (responsables de mise en œuvre et bénéficiaires clairs) ou est-elle plus diffuse et complexe (nombreux acteurs, nombreux bénéficiaires directs et indirects)? Les méthodes basées sur le recueil d'incidences sont particulièrement utiles pour récolter des preuves de changements très diverses, sur l'ensemble d'un secteur d'intervention, sans se limiter aux indicateurs d'effets ciblés par l'intervention.
- Quel est le temps disponible pour l'évaluation? Le recueil de récits de changement ou de preuves d'incidences peut nécessiter une présence sur le terrain de l'intervention, et un temps de collecte et d'analyse de données important. L'ERI est cependant une méthode qui permet d'aboutir à des constats très rapidement.

## Conclusion

Les méthodes d'évaluation des effets, qu'elles soient quantitatives ou qualitatives, sont une partie importante de la boîte à outils des évaluateurs et des évaluatrices. La diversité des méthodes d'évaluation des effets permet de répondre à la diversité des interventions et des missions d'évaluations.

Pour les évaluations d'effets les plus simples (théories de changement claires, peu de facteurs confondants pour des effets immédiats ou intermédiaires ou changements limités dans le temps), les devis implicites doivent constituer le premier réflexe, car ils sont faciles à mettre en œuvre, et sont souvent peu coûteux. Ils ont toutefois leurs limites, et leur validité interne peut être menacée par de nombreux biais.

Ainsi, pour des situations plus compliquées (sans pouvoir s'attaquer aux situations réellement complexes), on fera appel au devis expérimental et aux devis quasi expérimentaux lorsque la procédure expérimentale ne peut pas être mise en place. Ces méthodes quantitatives permettent d'écarter de

nombreuses explications concurrentes, et de calculer l'effet net d'une intervention en se servant d'un contrefactuel mesuré auprès d'un groupe de contrôle ou calculé de différentes manières. Mais ces méthodes reposent sur la prémisse que l'on a seulement besoin des résultats sur un (ou à la limite quelques) indicateur d'effet quantitatif. Or, face à des interventions complexes, le choix d'un indicateur d'effet quantitatif est délicat, voire impossible, car les effets sont diffus et moins prévisibles. C'est pourquoi la boîte à outils est enrichie par les méthodes qualitatives et mixtes qui répondent mieux aux besoins d'évaluation dans les situations de complexité.

La valeur technique d'une méthode d'évaluation des effets par rapport à une autre n'est, dès lors, pas suffisant pour en faire son option de prédilection. Le choix d'une méthode est de préférence déterminé par la plus-value que cette dernière peut apporter en termes de potentiels d'amélioration et de réponse à une préoccupation de gestion. On choisira dès lors les méthodes qui ne permettent pas simplement d'évaluer les effets, mais qui nous encouragent plutôt à réaliser une évaluation dans le but d'obtenir plus d'effets.

## Messages clés

### *Pourquoi évaluer les effets ?*

L'évaluation des effets est une question incontournable de la discipline. En effet, l'évaluateur ou l'évaluatrice est souvent le seul acteur qui dispose des méthodes permettant de pousser les investigations au plus loin de la chaîne des résultats. Là où un auditeur s'arrêtera à l'étude des extrants, l'évaluateur ou l'évaluatrice a les outils pour faire le lien entre une intervention et les changements qu'elle apporte à une population cible. Cette connaissance est le fondement même d'une gestion axée sur les résultats.

### *Qu'est-ce que nous apportent les méthodes d'évaluation des effets ?*

Les méthodes d'évaluation des effets sont différentes techniques permettant de faire le lien entre des changements et une intervention. L'évaluateur ou l'évaluatrice ne peut pas simplement observer une variation d'un indicateur cible et affirmer que la variation est due à l'intervention évaluée. Pour que son argumentaire évaluatif soit crédible, il doit s'appuyer sur une méthode qualitative, quantitative ou mixte qui lui permet d'écarter les autres facteurs d'influence sur la cible.

### *Faut-il privilégier certaines méthodes ?*

Il n'y a pas une méthode intrinsèquement plus pertinente ou robuste que les autres. Au-delà des clivages entre « quanti » et « quali », la bonne méthode est celle qui s'adapte le mieux à la situation évaluée : le calendrier et le budget de la mission, le besoin d'information des décideurs, la disponibilité et les possibilités de recueil des données, la complexité de la problématique, la diversité des parties prenantes, etc. La méthode la plus simple dans le contexte est souvent la bonne, et il ne faut pas hésiter à combiner et à hybrider les méthodologies pour les pousser plus loin (sans oublier de partager ses bons coups avec la communauté des évaluateurs et des évaluatrices).



## Pour aller plus loin

Disponible gratuitement en ligne, ce manuel de la Banque mondiale, traduit en français, est assez complet sur les méthodes quantitatives, et présente une diversité de cas réels, notamment issus de programmes de développement :

Gertler, P. J. *et al.* (2011). *L'évaluation d'impact en pratique*. Banque Mondiale.

Un document court, encore une fois traduit en français, présente non pas l'évaluation d'impact d'un point de vue pratique, mais plutôt selon la perspective de l'utilisateur et du demandeur d'évaluation :

Rogers, P. (2014). *Présentation de l'évaluation d'impact* (note méthodologique n° 1). Centre de recherche Innocenti de l'UNICEF. [https://www.unicef-irc.org/publications/pdf/brief\\_1\\_overview\\_fre.pdf](https://www.unicef-irc.org/publications/pdf/brief_1_overview_fre.pdf)

Un guide en ligne, publié par le Secrétariat du Conseil du Trésor du Canada sur l'intégration des approches axées sur la théorie d'intervention dans la pratique des évaluateurs et des évaluatrices du secteur public fédéral canadien :

Lien : <https://www.canada.ca/fr/secretariat-conseil-tresor/services/verifications-evaluations/evaluation-gouvernement-canada/approches-evaluation-axees-theorie-concepts-pratiques.html#fig4>

## Références bibliographiques

Anstett, E. (2018). *Méthodes statistiques et outils d'évaluation d'impacts* [matériel pédagogique]. École nationale d'administration publique (ENAP).

Anstett, E. (2019). *Programme international de formation en évaluation du développement (PIFED)* [matériel pédagogique]. École nationale d'administration publique (ENAP).

Baker, J. L. (2000). *Évaluation de l'impact des projets de développement sur la pauvreté : Manuel à l'attention des praticiens*. Banque mondiale.

Bamberger, M. (2012). *Introduction aux méthodes mixtes en évaluation d'impact – notes sur l'évaluation d'impact # 3*. InterAction.

Banque Mondiale. (2002). *Étude d'impact : Technique d'évaluation de programme d'aide active à l'emploi* (publication n° 2).

Centre d'excellence en évaluation. (2019). *Guide sur l'évaluation rapide de l'impact*. Gouvernement du Canada. <https://www.canada.ca/fr/secretariat-conseil-tresor/services/verifications-evaluations/centre-excellence-en-evaluation/guide-sur-levaluation-rapide-limpact.html>

Commission Européenne. (2013). *Concevoir et appliquer des évaluations d'impact contrefactuelles*. <https://op.europa.eu/en/publication-detail/-/publication/f879a9c1-4e50-4a7b-954c-9a88d1be369c/language-fr>

Copestake, J., Morsink, M. et Remnant, F. (2019). *Attributing Development Impact : the Qualitative Impact Protocol case book*. Practical Action Publishing.

Davies, R. et Dart, J. (2005). *La technique du changement le plus significatif CPS*. Monitoring and Evaluation News. <https://mande.co.uk/special-issues/most-significant-change-msc/>

- Dietz, T., Bélemvire, A., van der Geest, K., de Groot, D., Millar, D., Obeng, F., Pouw, N., Rijnveld, W. et Zaal, F. (2013). *PADev Manuel d'instruction*. African Studies Centre et Royal Tropical Institute. <http://www.padev.nl/guidebook.htm>
- Earl, S., Carden, F. et Smutylo, T. (2002). *La cartographie des incidences, intégrer l'apprentissage et la réflexion dans les programmes de développement*. Centre de recherches pour le développement international.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, R. B. et Vermeersch, M. J. C. (2011). *L'évaluation d'impact en pratique*. Banque mondiale.
- Hempel, K. et Fiala, N. (2011). *Mesurer la réussite des interventions auprès des jeunes : Guide pratique sur le suivi et l'évaluation*. Banque mondiale.
- IE4D Group. (2011). *Impact Evaluation for Development—Principles for Action*.
- Leeuw, F. et Vaessen, J. (2009). *Impact Evaluations and Development, Nonie Guidance on Impact Evaluation*. World Bank—Network of Networks for Impact Evaluation.
- Mayneris, F. (2009). *L'économie du développement à l'épreuve du terrain : entretien avec Esther Duflo*. La vie des idées.
- OCDE. (2002). *Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats*. The Development Assistance Committee (DAC) Working Party on Aid Evaluation.
- Secrétariat du Conseil du Trésor du Canada. (1998). *Méthodes d'évaluation des programmes : Mesure et attribution des résultats des programmes* (3<sup>e</sup> édition). <https://www.tbs-sct.gc.ca/cee/pubs/meth/pem-mep-fra.pdf>
- Wilson-Grau, R. (2019). *Outcome Harvesting : Principles, Steps, and Evaluation Applications*. Information Age Publishing.



# Les devis expérimentaux pour évaluer les mécanismes de lutte contre la pauvreté : défis et enjeux

*Deo-Gracias Houndolo*

## Introduction

Dans la littérature, il existe différentes définitions de la pauvreté et diverses façons de la mesurer. Elle est généralement considérée comme une question multidimensionnelle (Damoah et Akwei, 2017; Banerjee *et al.*, 2015; Bryceson, 2002) qui est abordée en termes absolus ou relatifs, et qui est perçue comme un manque de ressources pour satisfaire les besoins, ou l’incapacité à atteindre les capacités. Parfois, elle fait référence à la vulnérabilité, à l’inégalité, à l’exclusion, au fait d’être dans le besoin. Dans d’autres ouvrages, elle est abordée en termes d’état d’esprit et d’aspects non matériels.

Quel que soit l’angle de sa définition, la pauvreté est un fléau contre lequel toutes les communautés humaines tentent de lutter en adoptant des mécanismes différents. Au niveau mondial, la lutte contre la pauvreté est au cœur de plusieurs efforts de collaboration et est retenue comme l’objectif central et prioritaire des Objectifs du Millénaire pour le développement (OMD) en 2000, puis relayée encore dans les Objectifs de développement durable (ODD) en 2015. Depuis lors, plusieurs politiques, programmes et projets sont conçus et mis en œuvre pour réduire ou éradiquer la pauvreté. Quant aux méthodes d’évaluation de l’efficacité de ces mécanismes, elles sont variées, comme cela a été vu dans certains chapitres de ce manuel. Dans ce chapitre, nous partageons notre perspective sur les défis et les enjeux liés à l’application des devis expérimentaux (ou essais contrôlés randomisés – RCT) pour évaluer les mécanismes de luttés contre la pauvreté.

## Mécanismes de lutte contre la pauvreté et exemple d’application empirique d’un devis expérimental

L’élaboration d’une politique publique de lutte contre la pauvreté peut répondre à un objectif aussi spécifique que la réduction de la faim, l’amélioration de l’état nutritionnel des enfants de moins de cinq ans, l’augmentation du revenu agricole des producteurs, la réduction des dépenses de santé des ménages, ou encore l’amélioration des résultats scolaires des élèves, la réduction de la grossesse en milieu scolaire, l’amélioration de l’accès à l’eau potable en quantité journalière suffisante, etc.

À cet effet, les interventions suivantes peuvent être conçues pour lutter contre la pauvreté : le conseil agricole innovant pour les petits producteurs en milieu rural, les subventions de l’État aux intrants agricoles, l’éducation parentale à la diversification alimentaire dans les zones touchées par la

malnutrition, la gratuité des soins de santé de base pour les indigents, le transfert d'argent aux plus pauvres et aux personnes du troisième âge, la gratuité de l'accès à l'éducation primaire, la réalisation à grande échelle des infrastructures de base pour l'accès à l'eau potable, etc.

Ces différents mécanismes de lutte contre la pauvreté sont conçus et mis en œuvre à travers le monde depuis bien longtemps et plus intensément depuis les années 2000 avec les OMD et les ODD. Des ressources publiques y sont allouées, des institutions des Nations Unies<sup>1</sup>, et d'autres organismes régionaux et nationaux ainsi que les ONG et autres initiatives privées sont mis en place, toujours pour lutter contre la pauvreté qui, malheureusement, persiste toujours.

Il apparaît légitime de s'interroger sur l'efficacité et l'efficience des divers mécanismes de lutte contre la pauvreté, les raisons pour lesquelles ils ne permettent pas d'éradiquer la pauvreté, et d'explorer des pistes d'amélioration. Dans ce contexte, la question essentielle revient à se demander quel est l'impact des mécanismes de lutte contre la pauvreté. Les devis expérimentaux sont une des méthodes quantitatives qui permettent d'isoler de leur estimation d'impact les effets que les caractéristiques observables et non observables des bénéficiaires pourraient avoir sur les indicateurs de mesure de pauvreté.

Pour illustrer nos propos, prenons le cas concret d'un projet de Plan Foncier Rural qui vise à enregistrer les terres agricoles et à délivrer des certificats fonciers aux bénéficiaires. Un tel projet souhaite pallier un contexte rural où les terres agricoles ne sont pas régulièrement enregistrées par le gouvernement et où aucun propriétaire ne dispose de document légal justifiant le droit de propriété, ce qui ne favorise pas les investissements agricoles et qui contribue à terme à la pauvreté des producteurs vivant dans ces milieux. L'objectif du projet est que les bénéficiaires puissent obtenir une sécurisation suffisante de leurs terres et qu'ils soient à l'abri des conflits fonciers. Ils pourront, à moyen terme, faire des investissements durables sur leurs terres, avoir accès à du crédit bancaire parce qu'ils possèdent les documents de propriété, et participer plus efficacement à l'animation du marché foncier. À terme, le projet permettra d'améliorer les productivités agricoles, de connaître une augmentation nette du revenu agricole et une meilleure émancipation des femmes, qui sont autant d'éléments qui concourent tous à une réduction sensible de la pauvreté dans les communautés bénéficiaires. En vue d'évaluer l'impact d'une telle intervention, nous vous présentons comment l'assignation aléatoire des participants peut être réalisée<sup>2</sup> selon un devis expérimental.

La première étape de ce processus a été la définition de la population remplissant les conditions requises pour l'intervention. De ce point de vue, il était essentiel de définir le concept de pauvreté, de décrire les caractéristiques des personnes pauvres et de retenir les critères objectivement vérifiables qui font qu'une unité de la population est admissible ou non au projet. Notons que la pauvreté a été définie selon la perspective du manque de ressources, de la vulnérabilité, de l'inégalité et de l'exclusion. La deuxième étape a consisté à choisir la taille de la population, sa composition et sa distribution géographique. Dans ce cas de figure, ce sont les villages qui ont été ciblés comme unité d'assignation, tandis que les producteurs disposant de moins de 1,5 ha de terre agricole dans ces villages ont bénéficié de l'intervention et constituent l'unité d'observation. La sélection de l'échantillon, c'est-à-dire d'un sous-ensemble de cette population, a été déterminée selon un échantillonnage aléatoire avec possibilité

---

1. On pense à l'Organisation des Nations unies pour l'alimentation et l'agriculture (FAO), le Programme alimentaire mondial (PAM), le Fonds des Nations unies pour l'enfance (UNICEF) ou encore le Programme des Nations unies pour le développement (PNUD).

2. Les étapes de l'assignation aléatoire des participants ont été présentées dans le chapitre « Des méthodes d'évaluation des effets » de Anstett et Abdallah.

de remplacement. Concernant l'étape de la répartition géographique, les villages admissibles ont soumis leurs dossiers de candidature. Ensuite, les chefs de villages concernés se sont tous réunis dans un lieu public et ont procédé à un tirage aléatoire, une forme de loterie, pour déterminer quels villages seront bénéficiaires du projet pour constituer le groupe de traitement, et quels villages ne seront pas bénéficiaires, et qui vont par conséquent constituer le groupe contrôle ou témoin ou encore non bénéficiaires. Les producteurs de ces villages ont pris d'office le statut de traitement ou contrôle qui est celui de leur village. Par la suite, le projet a été mis en œuvre comme prévu, bien que certains producteurs qui habitent des villages non bénéficiaires aient parfois des terres agricoles admissibles dans des villages bénéficiaires, ou encore que des producteurs ayant des terres agricoles bien plus grandes que 1,5 ha déclarent une plus petite portion pour être bénéficiaires de l'intervention. Houngbédji (2015), Goldstein *et al.* (2016), Goldstein *et al.* (2018) et Wren *et al.* (2020) ont présenté les résultats de l'effet de ce projet.

En dépit des avantages et de l'intérêt offerts par les devis expérimentaux, leur adoption pour l'évaluation d'une intervention et leur mise en œuvre présentent divers défis et enjeux.

## **Défis et enjeux liés à l'application des devis expérimentaux pour l'évaluation des mécanismes de pauvreté**

Les devis expérimentaux sont caractérisés par divers types de défis dont la nature et l'intensité dépendent du niveau auquel on se situe dans la recherche.

Durant la phase de conception, il est courant que la randomisation se heurte à une résistance des politiques ou des populations elles-mêmes, soit du fait que certaines unités ne recevront pas l'intervention alors qu'a priori elles en ont besoin et la considèrent comme utile, nécessaire et parfois indispensable pour elles. Par ailleurs, d'autres évoquent des questions d'ordre éthique, et tentent de remettre en cause la démarche méthodologique qui est taxée de dénier à une frange de la population le droit d'accès à des biens et services publics disponibles en quantité suffisante et dont elles ont besoin. Ainsi, la faisabilité des défis expérimentaux est régulièrement remise en cause au premier abord. Pour y remédier, la considération des variantes de la randomisation fournit des solutions techniques pour ne pas abandonner la possibilité de conduire des devis expérimentaux.

La durée de mise en œuvre considérée comme relativement longue par les décideurs politiques et les commanditaires constitue un facteur qu'il faudra prendre en compte. Pour cela, il y a lieu d'examiner dans quelle mesure il est possible de produire des résultats intermédiaires, ou d'utiliser des indicateurs de résultats qui ne nécessitent pas un temps relativement long pour que des effets se manifestent sur eux.

Le coût des devis expérimentaux est parfois perçu comme trop élevé pour les budgets des projets, ce qui pourrait affecter la possibilité de les mettre en œuvre. À cet effet, il est recommandé de réaliser que les collectes de données constituent l'une des activités qui consomment une grande part des budgets. Ainsi, il faudrait procéder à une bonne détermination de la taille d'échantillon requise pour les évaluations de ce type pour éviter de gonfler le budget de l'évaluation sans que cela soit nécessaire. Autant que possible, l'examen des possibilités d'utilisation de données secondaires est une option qui aide également à réduire les coûts. Par ailleurs, l'anticipation du coût des évaluations dans les budgets de projets à leur conception est une solution bien efficace face au défi lié au coût des évaluations. Dans certaines circonstances, la recherche de partenaires financiers et techniques est également une réponse efficace aux contraintes budgétaires.

L'inexistence de compétences ou la faible capacité technique des parties prenantes des projets sur lesquels les évaluations de type expérimentales doivent se faire constituent un défi majeur à prendre en compte. En effet, il est courant de constater que certaines personnes s'opposent à l'adoption des devis expérimentaux parce qu'elles ont une compréhension erronée des exigences techniques, surtout en matière de constitution de groupe contrôle, de taille d'échantillon et de durée des expérimentations. Dans les situations où ces acteurs sont très influents, ils peuvent arrêter le processus d'évaluation. La solution recommandée face à cela est d'organiser des ateliers de formation à leur endroit tout au long du processus afin de renforcer leurs capacités. Ainsi, leurs craintes pourraient disparaître progressivement après les formations et les dialogues de sensibilisation sur le bien-fondé des devis expérimentaux. Ce qui permet d'en faire des alliés qui finissent par adhérer à la démarche méthodologique.

Dans les circonstances où il n'est pas possible de procéder à une randomisation simple, il faudra examiner des variantes de devis expérimental pour choisir la plus appropriée pour faire face aux défis. Ainsi, l'on pourrait considérer par exemple l'assignation aléatoire stratifiée, l'assignation aléatoire factorielle ou l'assignation aléatoire par phasage.

Dans un contexte de pauvreté ou d'urgence humanitaire, un devis expérimental par assignation aléatoire simple pourrait ne pas être faisable du fait de contraintes éthiques ou de défis pratiques de mise en œuvre liés à la mobilité des bénéficiaires par exemple. L'assignation aléatoire stratifiée pourrait, dans ces conditions, constituer une alternative crédible et faisable. Elle consiste à définir des strates ou niveaux d'intervention qui seront assignés aléatoirement à recevoir ou non le traitement. Il peut s'agir par exemple des écoles, des centres de santé des villages. Une fois l'assignation faite à l'intérieur de chaque strate, l'on définit un deuxième niveau d'assignation (par exemple les élèves d'une école, les patients d'un centre de santé, les ménages d'un village) dont les unités seront assignées ou non au traitement. Cette variante de devis expérimental a l'avantage d'offrir la possibilité d'une plus grande envergure en termes de zones de couverture et donne une meilleure variabilité en matière d'hétérogénéité des participants à l'évaluation.

L'assignation aléatoire factorielle consiste à disposer d'une intervention qui est constituée de plusieurs composantes. À cet effet, par exemple, il peut s'agir d'une intervention complexe qui a trois composantes : un appui matériel pour faciliter l'accès à l'eau de boisson; une subvention pour la construction de toilettes hygiéniques; un coffret offert pour assurer l'entretien et l'assainissement des infrastructures publiques telles que les marchés et les gares routières. Dans une démarche factorielle, chaque composante de l'intervention peut être considérée comme un traitement, et l'ensemble des composantes constituera l'intervention. Ainsi, au lieu de procéder comme dans une assignation aléatoire simple, on pourrait assigner les unités à quatre traitements, dont trois types de traitements, et une situation où les personnes concernées reçoivent toutes les trois traitements, soit l'intervention dans sa globalité. En adoptant cette variante du devis expérimental, cela permet de ne laisser aucune unité de la population sans un minimum d'appui. Ainsi, les tensions sociales et autres questions éthiques soulevées par le devis expérimental sont atténuées.

L'assignation aléatoire par phasage, qui est encore appelée phasage aléatoire, est une variante de devis expérimental qui consiste à mettre en œuvre la même intervention en faisant varier les groupes de bénéficiaires d'un cycle à l'autre. En fonction du nombre de cycles de mise en œuvre de l'intervention, les unités admissibles sont alors assignées de manière aléatoire pour être traitées durant l'une des



phases de l'intervention. Dans ces conditions, durant la première phase par exemple, les unités traitées vont constituer le groupe de traitement, et les unités qui ne sont pas encore traitées vont constituer le groupe de contrôle. Durant la phase suivante, une partie des unités du groupe, considérée comme contrôle lors de la phase initiale, sera traitée pour constituer à cette étape le groupe de traitement, et les unités qui ne sont toujours pas bénéficiaires du traitement vont continuer à être considérées comme le groupe contrôle. Ce processus se déroulera jusqu'à la fin de l'étude d'évaluation. Le phasage aléatoire aide à relever les défis liés à la nécessité d'apporter un appui à toutes les unités de la population admissible dans un contexte où il n'y a pas de problème majeur à différer l'obtention d'une intervention. Dans d'autres contextes, cette variante est adaptée lorsque les ressources humaines et matérielles sont limitées durant la première phase, ou encore lorsqu'il y a lieu de conduire une phase pilote pour produire des évidences qui aideront dans la mise en œuvre de l'intervention.

Lorsque l'équipe d'évaluation réussit à relever minutieusement ces défis en se basant sur la méconnaissance des risques encourus par les populations et les unités qui bénéficient de ces interventions, des défis d'ordre technique sont aussi à relever. Il s'agit notamment de la détermination de la taille d'échantillon requise pour détecter et mesurer un effet minimum lorsque ce dernier est effectivement causé par l'intervention. Les études à orientation quantitative font généralement face à ce défi, et sa résolution requiert l'implication de compétences techniques avérées.

Dans plusieurs situations, l'inexistence d'un cadre d'échantillonnage constitue un défi majeur dans la conception des devis expérimentaux. En effet, en l'absence de données sur les unités (ménages, entreprises, individus, centres de santé, écoles, groupements de producteurs, etc.), il est quasiment impossible de conduire l'évaluation. La réponse à une telle situation est de procéder à une énumération, à un recensement avant l'étude proprement dite, faire usage de données secondaires ou encore de procéder à un échantillonnage adapté à la circonstance.

Durant la mise en œuvre proprement dite, un défi courant est la localisation des unités ciblées. Dans ce cas de figure, l'utilisation des ressources fondées sur les systèmes d'information géographique (SIG), l'implication de personnes-ressources résidant dans le milieu et l'utilisation des informations précédemment collectées aident à relever le défi. La modification consciente des comportements et des élans naturels des bénéficiaires et/ou des non-bénéficiaires de l'intervention par eux-mêmes engendre respectivement l'effet Hawthorne, ou l'effet John Henry, qui entraînent une surestimation ou une sous-estimation de l'effet réellement attribuable à l'intervention. Dans ces conditions, l'utilisation de placebos, ou l'omission d'informer les participants de leur statut (traité ou contrôle) constituent des solutions pour faire face à ces défis.

En résumé, les défis peuvent être techniques, opérationnels, contextuels ou financiers dans la mise en œuvre d'un devis expérimental. Il y a lieu d'explorer les solutions les plus adaptées au nombre desquelles quelques-unes ont été proposées dans cet article. Outre les approches de randomisations, il faut bien noter que les approches quasi expérimentales et d'autres méthodes existent, et sont parfois tout aussi valides pour aider à la production de données probantes qui participent à la prise de décisions éclairées dans l'élaboration et la mise en œuvre des politiques publiques.



## Conclusion

Les devis expérimentaux sont utiles et nécessaires dans l'élaboration et la mise en œuvre des mécanismes de lutte contre la pauvreté, ainsi que leur reformulation, sans toutefois ignorer qu'ils ne sont pas une panacée. En effet, la validité des données probantes produites par les études d'évaluation constitue l'un des éléments clés de l'enjeu auquel font face les décideurs et les chercheurs. Toutefois, les biais qui émergent parfois durant la mise en œuvre des devis expérimentaux constituent de réelles menaces à leur intégrité comme cela a été vu dans le chapitre sur « Des méthodes d'évaluation des effets ». Dans un contexte global où les ressources sont limitées, les approches économiques en évaluation sont également intéressantes pour aborder l'enjeu économique et financier des interventions mises en œuvre pour lutter contre la pauvreté<sup>3</sup>.

À l'ère du numérique où les mégadonnées (*big data*) sont de plus en plus disponibles et accessibles, soutenues par les données satellitaires, également une source importante de données, et renseignent sur la pauvreté et le bien-être des populations, les devis expérimentaux s'enrichiraient à les intégrer dans le cadre de l'évaluation des interventions de lutte contre la pauvreté.

## Pour aller plus loin

Nous présentons trois références dont les perspectives sont complémentaires et couvrent aussi bien le champ théorique, intuitif qu'empirique des devis expérimentaux en tant que méthode d'évaluation des effets causés par une intervention. Ces références sont organisées sous forme d'entonnoir, en abordant d'abord les questions élargies avant de finir par les aspects spécifiques de mise en œuvre d'un exemple de devis expérimental.

**Fougère, D. (2010). Les méthodes économétriques d'évaluation. *Revue française des affaires sociales*, (1), 105-128.**

<https://www.cairn.info/revue-francaise-des-affaires-sociales-2010-1-page-105.htm>

Cet article présente le cadre méthodologique global de l'évaluation des politiques publiques en utilisant des approches quantitatives basées sur les statistiques et l'économétrie. Elle est une importante contribution au savoir qui fait la revue des principes fondamentaux des devis expérimentaux et non expérimentaux. L'auteur présente clairement le problème de biais de sélection et du défi méthodologique relatif à la causalité. Les modèles canonique (Rubin, 1974) et économétrique (Roy, 1951) y sont présentés. Il aborde également les évaluations en contexte non expérimental et couvre les méthodes quasi expérimentales telles que l'appariement, la double différence et la régression sur discontinuité dont les avantages et les limites sont spécifiquement abordés. Il s'agit d'une référence pour se faire une bonne idée globale et intuitive des méthodes économétriques pour estimer l'impact attribuable à des interventions de développement.

**Duflo, E. (2005). Évaluer l'impact des programmes d'aide au développement : le rôle des évaluations par assignation aléatoire. *Revue d'économie du développement*, 13(2), 185-226.**

<https://www.cairn.info/revue-d-economie-du-developpement-2005-2-page-185.htm>

---

3. Cf. les chapitres de cet ouvrage « Les approches économiques en évaluation : un aperçu » par Anstett et Ouédraogo ainsi que « L'approche de la Value for Money (VFM) de King » par Rey et Fortin.

Cet ouvrage fait l'apologie des évaluations basées sur les devis expérimentaux. Il présente en quoi les résultats issus de cette approche méthodologique sont valides et peuvent être utilisés dans les prises de décision en matière de politiques publiques, de mise en œuvre des projets et surtout comment les devis expérimentaux sont utiles dans le passage à l'échelle des projets, et les échanges d'expériences entre pays, entre régions par rapport à la conception et à la mise en œuvre des mécanismes de développement. Il constitue une bonne référence méthodologique soutenue par des exemples concrets issus de réalités pratiques. Contrairement à d'autres références, il situe la place des organisations internationales dans la promotion des devis expérimentaux, et également dans le financement de telles études. Cet article est un argumentaire en faveur des devis expérimentaux que nous recommandons, car il tire sa force de l'encrage technique et des illustrations pratiques.

**Houndolo, D. G., Hodonou, A., Sossou, S. et Hamidou Yacoubou, R. (2020). A Pathway to Adoption of Yield-Enhancing Agricultural Technologies among the Rural Poor : Evidence from a Randomized Control Trial in Benin. *Partnership for Economic Policy Working Paper*, (2020-15).**

[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3673487](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3673487)

Cet article est une illustration empirique du processus de mise en œuvre d'un devis expérimental. Il aborde la réalité du terrain, les défis et les enjeux, et illustre bien comment s'évalue rigoureusement un projet de développement, dans un contexte de pauvreté. La description de l'intervention qui y est faite illustre bien l'assignation aléatoire factorielle telle que décrite précédemment. Les auteurs ont fondé leurs analyses et discussions sur une approche économétrique étroitement liée à la théorie de changement de l'intervention. Les effets tels que mesurés et présentés montrent que les effets ont un itinéraire et ne sont pas fixes dans le temps. Ce qui indique le caractère dynamique des effets mesurés, même en utilisant les devis expérimentaux, et l'importance d'en tenir compte dans l'utilisation des évidences pour améliorer les politiques publiques de lutte contre la pauvreté.

## Références bibliographiques

- Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Shapiro, W., Thuysbaert, B. et Udry, C. (2015). A multifaceted program causes lasting progress for the very poor : Evidence from six countries. *Science*, 348(6236).
- Bryceson, D. F. (2002). The scramble in Africa : reorienting rural livelihoods. *World development*, 30(5), 725-739.
- Damoah, I. S. et Akwei, C. (2017). Government project failure in Ghana : a multidimensional approach. *International Journal of Managing Projects in Business*, 10(1), 32-59.
- Duflo, E. (2005). Évaluer l'impact des programmes d'aide au développement : le rôle des évaluations par assignation aléatoire. *Revue d'économie du développement*, 13(2), 185-226.
- Fougère, D. (2010). Les méthodes économétriques d'évaluation. *Revue française des affaires sociales*, (1), 105-128.
- Goldstein, M., Hounghbedji, K., Kondylis, F., O'Sullivan, M. et Selod, H. (2016). Formalizing rural land rights in West Africa : Early evidence from a randomized impact evaluation in Benin. *World Bank Policy Research Working Paper*, (7435). <http://hdl.handle.net/10986/22857>

- Goldstein, M., Hounghbedji, K., Kondylis, F., O'Sullivan, M. et Selod, H. (2018). Formalization without certification? Experimental evidence on property rights and investment. *Journal of Development Economics*, 132, 57-74. <https://doi.org/10.1016/j.jdeveco.2017.12.008>
- Houndolo, D. G., Hodonou, A., Sossou, S. et Hamidou Yacoubou, R. (2020). A Pathway to Adoption of Yield-Enhancing Agricultural Technologies among the Rural Poor : Evidence from a Randomized Control Trial in Benin. *Partnership for Economic Policy Working Paper*, (2020-15).
- Roy, A. D. (1951). Some Thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2), 135-146. <https://doi.org/10.1093/oxfordjournals.oep.a041827>
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701. <https://doi.org/10.1037/h0037350>
- Wren-Lewis, L., Becerra-Valbuena, L. et Hounghbedji, K. (2020). Formalizing land rights can reduce forest loss : Experimental evidence from Benin. *Science Advances*, 6(26).

# L'engagement des enfants et des jeunes dans les évaluations

*Michele Tarsilla*

## Introduction

Depuis une décennie, des mots tels que «équité», «diversité», «participation» et «autonomisation» font partie du lexique non seulement des bailleurs de fonds, mais aussi des gestionnaires des évaluations et des évaluateurs et évaluatrices proprement dits. Il n'est pas donc surprenant qu'un nombre croissant de professionnels de l'évaluation adoptent d'une part des outils de collecte adaptés de manière à mieux cerner la voix des plus marginalisés, et de l'autre, des procédés d'analyse des données qui privilégient la désagrégation selon le sexe, l'âge et le milieu de résidence (urbain/rural). Toutefois, tandis que les rapports de démarrage et les sections méthodologiques de beaucoup de rapports d'évaluation semblent confirmer l'adoption de telles pratiques évaluatives assez innovantes, un examen plus rigoureux des constats, des conclusions et des recommandations de nombreuses évaluations, ainsi qu'un suivi de proximité des processus évaluatifs qui y sont associés, révèle le manque d'un vrai engagement des enfants lors de la planification, de la conduite et de la dissémination des dites évaluations. Cette observation est encore plus troublante lorsque l'on considère que plus de 47 % de la population africaine est représentée par des jeunes de moins de 18 ans (UNICEF, 2017). Donc, leur participation aux différentes étapes du processus évaluatif s'impose afin de promouvoir non seulement la diversité des parties prenantes et la représentativité des conclusions des évaluations, mais aussi l'acceptabilité et l'utilité de recommandations correspondantes. Les lignes directrices de l'évaluation (y compris celles du *United Nations Evaluation Group* (UNEG)) rappellent que les outils de collecte et d'analyse ainsi que les échantillons utilisés lors des évaluations devraient être les plus représentatifs possible. Pourtant, le pourcentage des évaluations des interventions étant censées produire un impact sur la vie et le bien-être des enfants et des jeunes et employant des approches ou des méthodes centrées sur les bénéficiaires les plus jeunes, reste toujours très bas. Par exemple, l'engagement des enfants et des adolescents dans le développement des Termes de Référence (TdR), dans la revue des rapports provisoires et lors des réunions de restitutions des rapports finaux, demeure extrêmement rare de nos jours. Par ailleurs, même lorsqu'il y a une claire ambition d'engager les enfants dans les évaluations, le bassin d'évaluatrices et d'évaluateurs sur lesquels on pourrait s'appuyer pour mener ce type de mandat reste assez restreint, surtout à cause du manque de familiarité de la part des évaluateurs et évaluatrices avec les cadres théoriques et les méthodes évaluatives qui sont véritablement centrées sur les enfants.

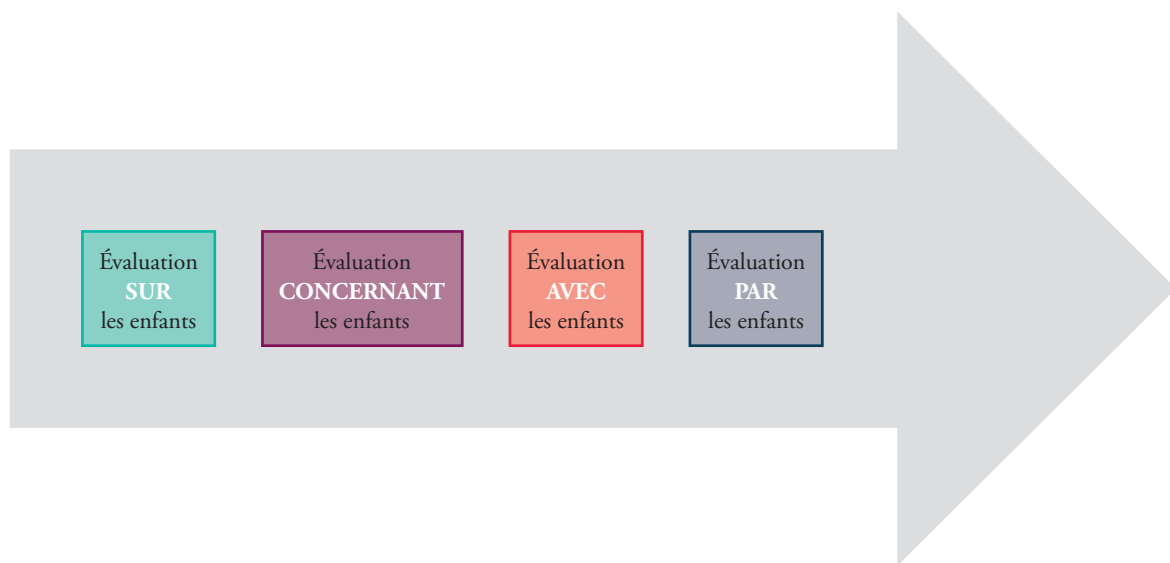
Dans le but de sortir de cette impasse et de stimuler une discussion qui puisse faire bouger les «cœurs et l'esprit» des évaluateurs et évaluatrices ainsi que des gestionnaires des évaluations, ce chapitre a pour objectif de présenter cinq stratégies clés ainsi que dix ressources (lignes directrices et manuels méthodologiques) qui pourraient aider les commanditaires et professionnels de l'évaluation à mieux impliquer les enfants («tout être humain de moins de 18 ans», selon la définition de la Convention Internationale des droits de l'enfant) dans les évaluations futures. Donc, en surmontant la pénurie

de connaissances et des pratiques exemplaires dans ce domaine, ce chapitre vise à renforcer la qualité et l'utilité des évaluations centrées sur les enfants dans le futur. Pour ce faire, ce chapitre transitoire sera bâti sur les conclusions de la Journée de Réflexion sur les Évaluations centrées sur les enfants, organisée par le Bureau du Fonds des Nations Unies pour l'enfance (UNICEF) pour l'Afrique de l'Ouest et du Centre (WCARO) le 4 décembre 2019. Plus spécifiquement, il familiarisera le lecteur avec les principes normatifs, les cadres théoriques de référence et les outils méthodologiques clés pour assurer une prise en compte plus adéquate des enfants tout au long du processus évaluatif.

Cette discussion semble encore plus pertinente à un moment comme celui que nous venons de traverser, où l'utilisation d'outils de collecte de données à distance (en raison des restrictions de voyage dues à la pandémie de COVID-19) a risqué – et continue de le faire – d'affecter la voix des enfants touchés par la crise sanitaire et socio-économique récente. À la fin du chapitre, les lecteurs pourront aussi bénéficier de deux petites surprises.

## **Stratégie n° 1: Essayez de « joindre le geste à la parole » et engagez pleinement les enfants dans vos évaluations.**

Dans ce but, dirigez-vous autant que possible vers la droite du spectre d'évaluation centrée sur l'enfant (voir ci-dessous).



**Figure 1**

Spectre sur l'engagement des enfants dans les évaluations.

Comme illustré à la gauche du spectre de l'évaluation, la plupart des évaluations sont menées « SUR » les enfants depuis des décennies, ce qui signifie que les enfants ont été traités comme des « objets d'évaluation » (en anglais *evaluand*), et donc évalués par le biais d'un examen documentaire ou d'une analyse de données secondaires. De même, en parcourant le spectre vers la droite, beaucoup d'autres évaluations sont menées « CONCERNANT » les enfants, ce qui signifie que les répondants adultes (par exemple, les parents lors des entretiens de ménages organisés pour évaluer un programme de transferts monétaires,

ou bien les enseignants participant à des groupes de discussion tenus pour estimer l'impact d'un programme scolaire, etc.) parlent des problèmes des enfants – ainsi que des effets de l'intervention sur la vie de plus jeunes – en leur nom. Ensuite, en continuant notre examen du spectre vers la droite, il apparaît qu'un nombre croissant d'évaluations récentes, même si très peu nombreuses, sont entreprises en collaboration « AVEC » des enfants, ce qui signifie que les enfants – une fois l'autorisation éthique reçue – ont échangé directement avec les évaluateurs et évaluatrices sur les questions les concernant. Enfin, quelques autres évaluations, même si elles sont extrêmement rares, sont menées « PAR » des enfants, qui collectent eux-mêmes des données primaires et qui les analysent de manière participative, sous la coordination d'un ou de plusieurs professionnels de l'évaluation.

## **Stratégie n° 2 : Il n'est pas nécessaire de travailler sur l'évaluation d'une intervention pour l'enfance afin d'engager les enfants dans un mandat d'évaluation.**

À cette fin, engager les enfants dans une évaluation n'est pas seulement « quelque chose de bien à faire » – c'est un acte nécessaire de responsabilité et de redevabilité. Une intégration plus systématique des opinions et des expériences des enfants dans les évaluations contribuera à la formulation de recommandations plus efficaces en vue d'améliorer durablement les interventions concernant l'entière d'un pays ou d'une région.

## **Stratégie n° 3 : Lors de la conduite d'une évaluation centrée sur les enfants, il faudra suivre l'Approche fondée sur les droits de l'homme (Groupe des Nations Unies pour le développement durable, s. d.) et définir les questions et rubriques correspondantes en se basant sur des documents normatifs globaux :**

À savoir la Convention relative aux droits de l'enfant (1989) et la Charte africaine des droits et du bien-être de l'enfant (Organisation de l'unité africaine, 1990), qui reconnaissent à la fois le droit des enfants à s'exprimer librement et le devoir des adultes d'écouter la voix des enfants.

## **Stratégie n° 4 : Il ne faut pas juste engager les enfants; il faut le faire de manière créative et innovante.**

Afin de rendre les interactions entre les équipes d'évaluation et les enfants plus efficaces et stimuler une participation accrue des plus jeunes, il sera nécessaire d'utiliser des outils de collecte de données participatifs et amusants comme le Cadre H, la Cartographie Corporelle (Save The Children, 2013, 2014a, 2014b)<sup>1</sup> et la Montgolfière.

1. Le Cadre H et la Cartographie corporelle sont présentés, parmi une panoplie d'outils pour l'évaluation de la participation des enfants, dans la *Trousse à outils du suivi et de l'évaluation de la participation des enfants*, par SaveTheChildren (2014a). Consultez, en particulier, le livret 5 : *Outils de suivi et d'évaluation de la participation des enfants*, [https://resource-centre-uploads.s3.amazonaws.com/uploads/me\\_toolkit\\_booklet\\_5\\_french.pdf](https://resource-centre-uploads.s3.amazonaws.com/uploads/me_toolkit_booklet_5_french.pdf)

## Le Cadre H

Cet outil peut être utilisé pour sonder les opinions et les sentiments des enfants par rapport aux forces (colonne de gauche) et aux faiblesses (colonne de droite) d'une intervention (parfois une activité de celle-ci) mise en œuvre dans leur communauté. Le nom de l'intervention ou de l'activité en question est normalement écrit par les enfants dans le petit espace rectangulaire au milieu (en haut). L'espace au milieu du cadre est utilisé par les enfants pour suggérer des améliorations concernant l'intervention ou l'activité évaluée.



**Figure 2**

Exemple de cadre H

Source : Reproduit de SaveTheChildren, 2013, p. 33, traduit librement par l'auteur.

## La cartographie corporelle

La **cartographie corporelle** est un outil participatif adapté pour donner aux enfants et aux jeunes la possibilité de partager librement leurs expériences après une urgence, comme la COVID-19. Si elle est correctement appliquée, la cartographie corporelle peut être un outil puissant pour comprendre les changements dans la vie des enfants « avant et après » un événement traumatique ou une intervention humanitaire.

Cet outil permet aux enfants d'exprimer leurs avis et de partager leurs expériences en matière de changements observés dans leur vie (ex. : pendant la COVID-19, par rapport à leur situation avant la pandémie). Des questions ciblées de suivi sont posées, notamment pour savoir dans quelle mesure le programme faisant l'objet de l'évaluation a influencé ces changements. L'image du corps (et les parties du corps qui y figurent)<sup>2</sup> est utilisée comme point de mire pour explorer et enregistrer les opinions des participants sur les changements apportés en matière de soins de santé, de santé sexuelle et reproductive, de mariage précoce, de santé mentale, d'hygiène (lavage des mains, troussees sanitaires), d'utilisation de masques, d'éducation, de responsabilités domestiques, etc. Cela peut être fait sur des petits papiers repositionnables ou des feuilles de papier qui peuvent être collés ou affichés sur l'image du corps (voir encadré 1).

2. L'image corporelle peut être dessinée à l'aide d'une grande feuille de papier, qui est produite en collant de grandes feuilles de papier pour tableau de conférence sur le sol, dans une pièce sécurisée. L'une des personnes représentées est ensuite invitée à se porter volontaire pour s'allonger sur les feuilles afin de faire dessiner sa silhouette pour créer une grande carte de son corps (comme le dessin d'une ombre).



### **Encadré 1 – Questions clés pour les enfants pour chacune des parties du corps\***

La tête : changements dans les connaissances des enfants, dans leur façon de penser et dans la façon dont les adultes pensent à eux.

- Comment fait-on pour éviter d'être contaminé par la COVID-19?
- Depuis quand connaissez-vous ces moyens?
- Par quels canaux avez-vous appris cela?
- Parmi ces moyens de prévention, lesquels sont plus faciles à mettre en pratique pour les enfants? Lesquels sont plus difficiles?

Les yeux : changements dans la façon dont les enfants se voient eux-mêmes et voient leur famille, leur communauté, leur école ainsi que le changement dans la façon dont les adultes les voient.

- Depuis le début de la COVID-19, quels changements avez-vous observés dans les relations que vous entretenez avec votre famille?
- Quels changements avez-vous remarqués dans la manière dont vous vous voyez? La façon dont vous regardez votre famille et les autres?

Les oreilles : changements dans votre façon d'écouter et d'être écouté par les autres.

- Quels sont les changements dans votre façon d'écouter depuis l'apparition de la COVID-19?
- Quels sont les changements dans la façon dont les autres vous écoutent depuis l'apparition de la COVID-19?

La bouche : changements dans la façon dont les enfants parlent, et que les adultes parlent d'eux et avec eux.

- Depuis l'apparition de la COVID-19, quels sont les changements intervenus dans votre façon de parler entre vous?
- La façon dont vous parlez avec les autres (adultes)?
- La façon dont les adultes parlent avec les enfants?

Les épaules : changements dans les responsabilités des enfants (filles et garçons).

- Depuis que la COVID-19 existe, quels sont les changements que vous avez constatés dans vos responsabilités (dans vos devoirs, vos rôles)?

Le cœur : changements dans la façon de se sentir et dans les attitudes des enfants envers les autres et des autres envers eux.

- Depuis le début de la COVID-19, comment vous êtes-vous senti(e)s?

*Relance* : Quels sentiments ressentez-vous depuis l'apparition de la COVID-19?

- Depuis l'apparition de la COVID-19, quels sont les changements dans la façon dont les autres se sentent vis-à-vis de vous ?

L'estomac : changements dans ce que les enfants mangent.

- Depuis l'apparition de la COVID-19, quels sont les changements intervenus dans ce que vous mangez ? Et dans votre façon de manger ?

Les mains et les bras : changements dans ce que les enfants font comme activités et la façon dont les adultes les traitent.

- Depuis que la COVID-19 existe, quels sont les changements que vous avez constatés dans vos activités quotidiennes ?
- Ces changements sont-ils restés comme tels jusqu'à maintenant ? Si non, à quels moments y a-t-il eu de nouveaux changements ?
- Quels changements avez-vous constatés à l'école ? Depuis quand ?
- Quels changements avez-vous constatés dans votre façon de travailler à l'école ? Depuis quand ?
- Quels changements avez-vous constatés dans la façon de travailler des autres enfants à l'école ? Depuis quand ?
- Connaissez-vous des enfants qui doivent mendier ou travailler hors de chez eux à cause de la COVID-19 parce que leur famille n'a plus de revenus ?
  - Si oui, combien d'enfants connaissez-vous dans cette situation ?
- Quels sont les changements intervenus dans la façon dont les adultes vous traitent ?

Les pieds et les jambes : changements dans leurs déplacements.

- Depuis l'apparition de la COVID-19, quels sont les changements intervenus dans vos déplacements ?

\*Dans le cadre de cet exemple, la COVID-19 a été utilisée à plusieurs reprises. Toutefois, l'outil reste applicable à plusieurs contextes. Donc, le terme « COVID-19 » pourrait être tout simplement remplacé par un autre (ex. « depuis l'ouverture du nouvel espace pour les enfants »; « depuis le tremblement de terre affectant votre communauté », etc.).

Contrairement aux autres outils participatifs où les deux sexes peuvent être regroupés (si nécessaire) pour le même exercice, il faut rappeler que l'exercice de cartographie corporelle doit être facilité par un agent de collecte du même sexe que celui des enfants qui prennent part à l'activité. Il est en outre prévu qu'un homme facilite l'exercice pour les garçons, et une femme pour les filles. Par conséquent, il est de la responsabilité du partenaire d'exécution sur le terrain de prendre en compte l'équilibre entre les sexes et de considérer l'expérience lors du recrutement des agents communautaires pour cet exercice participatif au niveau de la communauté. De plus, lorsque les exercices de cartographie corporelle sont effectués simultanément dans la même communauté, il faut envisager de faire appel

à des membres de la communauté pour aider au processus de facilitation. Le matériel nécessaire pour mettre en pratique cet outil inclut une grande feuille de papier, un grand tableau, des marqueurs, des autocollants de type « Post-it », des feuilles A4, et une colle/gomme.

### La montgolfière

L'outil de la **montgolfière** peut être adapté pour générer les besoins/problèmes initiaux identifiés par les enfants au sein de leur communauté, les types d'interventions et les meilleures pratiques qui ont été introduits pour répondre aux besoins/problèmes, et les facteurs qui empêchent la réalisation de l'intervention. Dans le contexte d'une évaluation participative engageant les enfants au sein d'une même communauté, l'outil de la montgolfière pourrait bien compléter les conclusions tirées lors de la cartographie corporelle mentionnée dans la section antérieure. La montgolfière est composée d'une nacelle, d'un ballon et d'une corde qui relie le ballon à la nacelle. Une montgolfière peut représenter un thème (par exemple la santé des enfants). Les participantes et participants sont invités à insérer dans la nacelle les principaux problèmes/défis/questions/besoins en matière de services de garde d'enfants. Ils sont ensuite priés de décrire les interventions et les meilleures pratiques qui ont été réalisées pour résoudre lesdits problèmes et de les représenter sous forme de corde liant la nacelle au ballon de la montgolfière. Il leur est ensuite demandé de visualiser les facteurs de résistance au changement qui ont fait descendre le ballon et de les représenter sous forme de fardeaux empêchant la montgolfière de s'envoler (ceci permet de documenter la raison pour laquelle le changement n'a pas eu lieu). Enfin, il leur est demandé de faire des suggestions pour éliminer ces facteurs inhibiteurs.



**Figure 3**

#### Exemple de montgolfière.

Source : Ballon à Air Chaud, Montgolfière, Festival de Montgolfières. (s. d.).  
 Freepng.fr. <https://www.freepng.fr/png-zib78b/>

Pour ce faire, les activités de collecte et d'analyse des données sur le terrain garderont un aspect ludique et léger, et éviteront de répliquer les dynamiques et « rituels » parfois rigides qui sont associés aux entretiens et aux groupes de discussion tenus avec les adultes. Cependant, pour que ces outils soient efficaces et transformationnels, il faut s'assurer qu'ils sont adaptés à la tranche d'âge spécifique des enfants ou des jeunes (par exemple : 6 à 10 ans; 11 à 14 ans; 15 à 18 ans, 19 à 24 ans) avec lesquels se déroulera le travail sur le terrain. Un autre outil qui mérite d'être exploré afin de mieux s'engager avec les enfants au cours des évaluations est la Photo Voix (celle-ci est traitée dans le chapitre «Une panoplie d'outils de collecte de données pour les évaluations»).

## **Stratégie n° 5 : Il est impératif de gagner de l'expérience du terrain en matière d'évaluation centrée sur les enfants, afin d'éviter le risque de « faire du mal » lors de l'utilisation de ces outils innovants auprès des enfants.**

Cela pourrait impliquer de : (i) mener des activités d'évaluation à titre bénévole avec des groupes de jeunes de sa propre communauté, toujours sous la supervision d'un consultant plus expérimenté; (ii) discuter avec d'autres évaluateurs et évaluatrices de la littérature spécialisée sur l'évaluation centrée sur les enfants<sup>3</sup>; (iii) consulter des sites Web spécialisés<sup>4</sup>; (iv) étudier la Recherche éthique impliquant des enfants (Graham *et al.*, 2013).

### **Conclusion**

Ce chapitre a tenté d'aller au-delà de toute généralisation et particularisation des enfants africains, trop souvent décrits et représentés comme sans espoir, faibles et sans aucune capacité de réagir. Au contraire, le but a été celui de transmettre aux lecteurs un message plus positif et inspirant : les enfants ne sont pas le problème, mais font plutôt partie de la solution aux défis de notre société contemporaine, en Afrique et ailleurs. L'évaluation se révèle donc un véhicule assez efficace pour contribuer à la valorisation des enfants et à la promotion d'une place plus importante dans la formulation des préoccupations politiques de beaucoup de pays. Toutefois, comme les enfants ne représentent pas un univers monolithique, ce chapitre a aussi mis en exergue que les approches et outils des évaluations devront être mieux contextualisés, en s'adaptant, par exemple, à l'âge et au niveau des capacités des enfants et des jeunes à s'engager lors du travail évaluatif sur le terrain.

### **Pour aller plus loin**

Comme anticipé au début de ce chapitre, les lecteurs pourront trouver encore plus d'informations et d'inspiration sur la façon de s'engager davantage avec les enfants dans les évaluations, en visionnant deux animations en 3D développées par l'UNICEF (voir les liens ci-dessous). Dans le but de promouvoir une plus grande prise de conscience de cette thématique au sein de la communauté de l'évaluation et d'augmenter l'inclusion de processus évaluatifs à l'échelle globale, les lecteurs seront encouragés à partager ces deux animations avec leurs pairs et avec leur réseau.

1. UNICEF Africa. (2020). *Inclure la voix des enfants dans la planification d'une évaluation* [vidéo]. YouTube. <https://www.youtube.com/watch?v=pl-YTrSrblM&list=PLRVDB06qgivs8FjXJ8tbiwuhgskg2-g7R&index=1>
2. UNICEF Africa. (2020). *Comment mieux vous engager avec les enfants dans vos évaluations lors de la collecte de données* [vidéo]. YouTube. [https://www.youtube.com/watch?v=LUdp\\_u0yorks&list=PLRVDB06qgivs8FjXJ8tbiwuhgskg2-g7R&index=2](https://www.youtube.com/watch?v=LUdp_u0yorks&list=PLRVDB06qgivs8FjXJ8tbiwuhgskg2-g7R&index=2)

3. Consulter le lien suivant : <https://www.unicef-irc.org/journal-articles/>

4. Consulter le lien suivant : [https://www.betterevaluation.org/en/themes/evaluation\\_and\\_children](https://www.betterevaluation.org/en/themes/evaluation_and_children)

## Références bibliographiques

- Convention relative aux droits de l'enfant*, New York, 20 novembre 1989, Nations Unies, Recueil des Traités, vol. 1577, p. 3. <https://www.unicef.fr/sites/default/files/convention-des-droits-de-lenfant.pdf>
- Graham, A., Powell, M., Taylor, N., Anderson, D. et Fitzgerald, R. (2013). *Recherche éthique impliquant des enfants*. UNICEF. [https://childethics.com/wp-content/uploads/2015/04/ERIC-compendium-FR\\_LR.pdf](https://childethics.com/wp-content/uploads/2015/04/ERIC-compendium-FR_LR.pdf)
- Groupe des Nations Unies pour le développement durable. (s. d.). *Approche fondée sur les droits de l'homme*. Nations Unies. <https://unsdg.un.org/fr/2030-agenda/universal-values/human-rights-based-approach>
- Organisation de l'unité africaine. (1990). *Charte Africaine des Droits et du Bien-être de l'Enfant*. <https://au.int/fr/treaties/charte-africaine-des-droits-et-du-bien-etre-de-lenfant>
- Rossion, D. (2015). *La participation des enfants et des jeunes dans l'évaluation des politiques publiques : Cadre théorique et analyse de cas* (document de travail n° 4). Observatoire de l'Enfance, de la Jeunesse et de l'Aide à la Jeunesse. [https://www.brudoc.be/opac\\_css/index.php?lvl=publisher\\_see&id=764](https://www.brudoc.be/opac_css/index.php?lvl=publisher_see&id=764)
- Save The Children. (2013). *Guidelines for Children's Participation in Humanitarian Programming*. <https://resourcecentre.savethechildren.net/document/guidelines-childrens-participation-humanitarian-programming/>
- Save The Children. (2014a). *Trousse à outils du suivi et de l'évaluation de la participation des enfants*. [https://resourcecentre.savethechildren.net/pdf/me\\_toolkit\\_booklet\\_1\\_french.pdf/](https://resourcecentre.savethechildren.net/pdf/me_toolkit_booklet_1_french.pdf/)
- Save The Children. (2014b). *Trousse à outils du suivi et de l'évaluation de la participation des enfants. Livret 5. Outil de suivi et d'évaluation de la participation des enfants*. [https://resource-centre-uploads.s3.amazonaws.com/uploads/me\\_toolkit\\_booklet\\_5\\_french.pdf](https://resource-centre-uploads.s3.amazonaws.com/uploads/me_toolkit_booklet_5_french.pdf)
- Shung King, M., September, R., Okatcha, F.M. et Cardoso, C. (2009). *Child research in Africa*. CODESRIA. <https://www.codesria.org/spip.php?article1314>
- Skelton, T. (2008). Research with children and young people : exploring the tensions between ethics, competence and participation. *Children's Geographies*, 6(1), 21-36. <https://doi.org/10.1080/14733280701791876>
- UNICEF. (2017). *Génération 2030. Afrique 2.0. Favoriser les investissements dans l'enfance pour bénéficier du dividende démographique*. [https://www.unicef.org/media/48691/file/Generation\\_2030\\_Africa\\_2.0\\_FR.pdf](https://www.unicef.org/media/48691/file/Generation_2030_Africa_2.0_FR.pdf)



# De l'usage des technologies de l'information et de la communication en évaluation

*Tigawendé Prosper Kaboré*

*Faya François Bourouno*

## Introduction

L'exercice de l'évaluation est souvent confronté à un certain nombre de défis et de contraintes. Les bailleurs et les autres acteurs du développement deviennent cependant de plus en plus exigeants sur la redevabilité, l'inclusion et l'impact réel des interventions sur les communautés. Les professionnels de l'évaluation, qu'ils soient praticiens ou commanditaires, sont régulièrement confrontés à de multiples contraintes. Dans les pays en développement, des contraintes spécifiques sont présentes, par exemple, les problèmes d'infrastructures compliquant l'accès à certaines zones défavorisées et rendant ainsi la collecte des données difficile ou impossible, le transport des fiches de collecte (en version papier) s'avère parfois fastidieux et peut rendre long le processus de traitement des données évaluatives (surtout les mandats qui exigent des évidences sur les ménages pauvres), les problèmes d'insécurité dans certaines zones d'intervention à cause des conflits armés, etc. Au fil du temps, les évaluateurs et évaluatrices se tournent vers les Technologies de l'Information et de la Communication (TIC) en tant que solution pratique aux multiples contraintes qu'ils rencontrent sur le terrain.

Ce chapitre présente des solutions technologiques couramment utilisées par les évaluateurs et évaluatrices pour résoudre certaines contraintes et minimiser les risques de biais, d'insécurité, d'exclusion et de représentativité. Il présentera brièvement également l'utilisation des nouvelles technologies pour le traitement des données qualitatives et quantitatives ainsi que pour la diffusion des résultats d'évaluation. Enfin, il présentera les limites liées à l'utilisation des TIC dans une évaluation.

## Contexte et définitions des TIC en évaluation

La collecte des données demeure une phase importante dans la démarche d'évaluation d'une intervention. De nombreux défis tant opérationnels que méthodologiques ayant un impact sur la qualité et l'opportunité des exercices d'évaluation ont été constatés par les praticiens. Les évaluatrices et évaluateurs sont confrontés régulièrement aux contraintes liées au temps, à la qualité, à l'exhaustivité et à la disponibilité des données, à la viabilité et à la sécurité des données. Dans cette perspective, les évaluateurs et évaluatrices recherchent des systèmes plus rigoureux, et en même temps plus souples, pour suivre et évaluer les interventions (Raftree et Bamberger, 2014) afin de contribuer à relever ces défis et à surmonter les problèmes méthodologiques.



Ainsi, les évaluatrices et les évaluateurs se sont tournés vers les TIC en tant que solution pratique aux multiples contraintes qu'ils rencontrent sur le terrain. Parmi les avantages des TIC en évaluation, on peut citer la diversification des options de collecte de données pour les équipes d'évaluation, la réduction des coûts, le gain de temps et la promotion des différentes formes de communication avec divers groupes de parties prenantes (De Miranda, 2019; Meador, 2020). Par exemple, une simple stratégie de messagerie texte, développée par Extensio au Mexique, a permis aux petits producteurs ruraux de recevoir des messages sur leurs téléphones portables avec des informations en temps réel sur les prix du marché, les conditions météorologiques et les ravageurs de leur région. Cela leur a permis de prendre de meilleures décisions concernant leurs cultures et d'augmenter leurs revenus (De Miranda, 2019).

Dans ce contexte, qu'est-ce qu'on entend par TIC, surtout en évaluation? Les TIC comprennent le matériel (téléphones mobiles, tablettes, capteurs, ordinateurs portables et de bureau), les données qui peuvent être collectées, téléchargées sur ces appareils ou accessibles par Internet, et les logiciels qui permettent l'analyse des données sur ces appareils (Meador, 2020). Pour cet expert en gestion des données, les TIC «font référence à l'ordinateurs (plus les logiciels), les smartphones, les objets connectés, les drones, les satellites ... et ces outils peuvent intervenir à plusieurs niveaux, lors de la conception des outils de collecte, de la collecte proprement dite, du traitement, l'analyse et la visualisation» (Kaboré, 2021). D'autres spécialistes estiment que les TIC concernent «la télévision, la radio, le téléphone, GPS "*Global Positioning System*", plateforme de collecte de données en ligne, appareils photo, site Web, page Facebook, Twitter, groupe WhatsApp, etc.» (Kaboré, 2021). Il convient également de noter que certains outils mentionnés ci-dessous peuvent être utilisés à plusieurs étapes d'un projet d'évaluation, et ne sont donc pas limités à une étape spécifique de l'évaluation (Meador, 2020).

Les différentes solutions pourraient être classées dans de nombreuses catégories différentes, car chacune d'entre elles a une série d'objectifs et d'utilisations. Il existe également une interconnexion naturelle entre de nombreux outils et applications des nouvelles technologies (Meador, 2020). La section suivante détaillera les différentes catégories et leur utilisation pour le traitement des données qualitatives et quantitatives.

## Utilisation des nouvelles technologies pour la collecte et le traitement des données qualitatives et quantitatives

Les définitions des TIC se concentrent sur l'utilisation de la technologie pour transmettre l'information, mais certaines définitions traitent également du processus (transfert et collecte de l'information), de l'utilisation (médias sociaux et analyse) et de l'objectif (augmentation de la profondeur et de l'étendue des communications). Les lignes suivantes vont détailler les différentes solutions ou outils des TIC qui permettent de faciliter la collecte et le traitement des données quantitatives et qualitatives.

### Les technologies mobiles

Les technologies mobiles permettent un accès direct aux bénéficiaires, aux participants et aux citoyens, et cet accès se fait généralement par le biais des téléphones mobiles. Ce matériel permet le transfert d'informations par le biais de radiofréquences (Corlazzoli, 2014). Il s'agit également, mais pas exclusivement, de l'utilisation de téléphones par SMS pour les enquêtes et les sondages simples; de tablettes et de téléphones intelligents pour les questionnaires plus longs et approfondis; de stylos numériques pour capturer des notes manuscrites et les stocker sous forme numérique. L'utilisation des technologies

mobiles a le potentiel d'améliorer les processus de collecte de données d'évaluation de plusieurs façons. Elles permettent notamment de produire des quantités de données beaucoup plus importantes, et ce rapidement et de manière plus fiable (Meador, 2020). Par exemple, VOTO Mobile a utilisé des SMS et des réponses vocales interactives pour atteindre des communautés qui seraient autrement trop éloignées pour être contactées en personne, ce qui a permis de collecter des données auprès de millions de personnes dans des dizaines de pays (Bamberger *et al.*, 2016; Meador, 2020). De plus, elles peuvent permettre la collecte de données dans des zones qui, autrement, ne seraient pas accessibles aux évaluateurs et évaluatrices en raison de problèmes de sécurité et d'autres difficultés (Meador, 2020). Par exemple, à l'aide d'appareils mobiles, d'appareils photo avec GPS et de Skype, Save the Children a pu collecter des données dans une région restreinte de Somalie en 2015 (Humanitarian Practice Network, 2016; IFAD, 2017). Enfin, à un niveau plus micro, des tablettes et des photographies ont été utilisées pour construire des indicateurs. La méthodologie *Poverty Spotlight* de la *Fundacion Paraguaya* consiste à travailler avec les ménages pour autoévaluer leur sentiment de pauvreté multi-dimensionnelle à travers 50 indicateurs, en utilisant une série de photographies présentées sur une tablette (Burt, 2013).

### **Systèmes d'information géographique (SIG)**

Selon le *National Center for Geographic Information Analysis* (NCGIA), les systèmes d'information géographique (SIG) sont « du matériel, des logiciels et des procédures conçus pour soutenir la saisie, la gestion, la manipulation, l'analyse, la modélisation et l'affichage de données à référence spatiale pour résoudre des problèmes complexes de planification et de gestion » (Griffiths, 2012; Meador, 2020). Leur principal avantage réside dans la cartographie des « données contextuelles de manière visuelle, permettant aux utilisateurs de détecter des modèles dans un programme et son environnement » (Azzam et Robinson, 2013). Ils peuvent également être particulièrement utiles pour évaluer l'impact sur l'environnement en raison de leur capacité à saisir une vue d'ensemble d'une région spécifique (IFAD, 2017). Par exemple, les SIG peuvent être utilisés pour identifier les sites optimaux pour la réalisation de l'intervention; dans le cadre du suivi, pour identifier l'efficacité éventuelle de l'intervention; ou pour améliorer la qualité des données nécessaires à la mesure de la performance (Azzam et Jacobson, 2015). En fin, les SIG, à travers les téléphones portables équipés de GPS (Azzam et Robinson, 2013), et la technologie des satellites, permettent désormais de localiser et de suivre les populations qui migrent en raison de la sécheresse, des catastrophes naturelles et de la guerre (Bamberger *et al.*, 2016).

### **Les médias sociaux**

Les médias sociaux sont définis comme des « formes de communication électronique (telles que les sites Web de réseaux sociaux) par lesquelles les utilisateurs créent des communautés en ligne pour partager des informations, des idées, des messages personnels et d'autres contenus » (Social Media, 2020). À l'heure actuelle, Facebook (Meta), Twitter, YouTube, WhatsApp et Tik Tok sont des exemples des médias sociaux les plus utilisés, avec des millions de visites quotidiennes uniques dans le monde entier (Meador, 2020). Ainsi, les médias sociaux peuvent être utilisés pour partager des informations avec un large public, pour susciter une interaction avec le public et des réactions de sa part, et pour créer un dialogue entre les individus (Corlazzoli, 2014). Le programme de consolidation de la paix souligne le potentiel de l'analyse des médias sociaux dans la collecte d'informations en temps quasi réel et la mesure du sentiment dans les zones de conflit (Corlazzoli, 2014). Néanmoins, un certain

nombre de défis subsistent, tels que le biais de sélection, la fiabilité des informations, la sécurité et la confidentialité des utilisateurs (Corlazzoli, 2014). L'analyse des données des médias sociaux peut également être complexe, nécessitant de nouvelles aptitudes et compétences qui ne sont pas actuellement discutées dans le domaine de l'évaluation (Meador, 2020).

### **Mégadonnées « *Big Data* »**

Les mégadonnées « *Big Data* » désignent les quantités massives de données qui sont désormais générées quotidiennement dans le cadre de l'informatisation croissante des systèmes et des dossiers par les gouvernements et les entreprises (Corlazzoli, 2014; Gobble, 2013). Bien qu'elles soient largement sous-utilisées en évaluation, les données massives possèdent un fort potentiel pour le domaine (Meador, 2020; Petersson *et al.*, 2017a). Par exemple, qu'il s'agisse d'infrastructures de soins de santé massives, de schémas de circulation sur des routes très fréquentées ou d'achats effectués sur des plateformes d'achat en ligne, des milliards de bits de données sont enregistrés et stockés quotidiennement (Lefebvre-Nare *et al.*, 2017; Meador, 2020). Ces données massives pourraient être analysées pour répondre à des questions d'évaluation sur les changements de comportement des populations cibles, sur l'amélioration des services publics, etc. (Meador, 2020).

Contrairement aux formes traditionnelles de collecte de données telles que les enquêtes et les entretiens, l'analyse des données massives peut examiner plusieurs sources de données simultanément et le faire en temps réel (Mazzeo Rinaldi *et al.*, 2017). Par exemple, l'initiative *Global Pulse* des Nations Unies vise à utiliser les données massives pour prédire les taux de pauvreté dans les communautés (Raftree et Bamberger, 2014). Cela implique que les futures interventions administrées par les Nations Unies et d'autres organisations du secteur social pourront s'appuyer sur les données produites par les projets de mégadonnées (Meador, 2020).

En résumé, les TIC influencent le travail de l'évaluation. La pandémie de COVID-19 demeure un exemple puissant, dans la mesure où les évaluateurs et évaluatrices ne pouvaient pas facilement se déplacer sur le terrain. Plusieurs outils permettent de surmonter ces défis.

## **Utilisation des TIC pour la diffusion des résultats d'évaluation**

En même temps que certaines solutions facilitent la collecte et le traitement des données, les TIC permettent également de diffuser les résultats de l'évaluation à travers la visualisation. Cette dernière consiste à utiliser des données qualitatives ou quantitatives pour générer des images qui peuvent ensuite être utilisées à des fins « d'exploration, d'examen et de communication » (Evergreen et Metzner, 2013). Il existe une pression croissante pour que les rapports soient courts, concis et pertinents. Les outils de visualisation des données y contribuent (Corlazzoli, 2014). Pour Linda Raftree, ils favorisent une meilleure utilisation des données pour une prise de décision plus éclairée sur le cadrage et le raffinement du type de données à collecter (2014). Les cartes, les graphiques et les tableaux, basés sur les TIC, facilitent l'analyse des données collectées et leur partage avec la communauté (Raftree, 2014).

Dans le domaine de l'évaluation, « la visualisation des données suscite un intérêt sans cesse croissant. Cela peut être dû à une réforme en cours dans laquelle les praticiens et les universitaires ont commencé à adapter les techniques de visualisation des données pour ajouter de la clarté, de la valeur esthétique et de la pertinence aux résultats de l'évaluation dans un effort pour améliorer la consommation des résultats de l'évaluation » (Meador, 2020).

La visualisation des données s'obtient généralement à l'aide de sites Web et de logiciels de conception. Il s'agit par exemple, des infographies, des tableaux de bord ou de la vidéo.

- **Les infographies** sont des ensembles d'images, de diagrammes et de textes brefs qui permettent de présenter un sujet de manière facile à comprendre (Nediger, 2019). Ils tiennent en une page (Web) en décrivant un seul sujet à l'aide de plusieurs éléments, notamment des graphiques, des diagrammes et des illustrations (Corlazzoli, 2014). Ils sont de plus en plus populaires.
- **Les tableaux de bord** sont une interface conçue pour présenter les informations de plusieurs indicateurs, généralement liés à un sujet ou à un concept sous-jacent commun, d'une manière facilement compréhensible sur un seul écran; les tableaux de bord sont une forme hautement interactive de visualisation des données (Corlazzoli, 2014).
- **La vidéo** peut être utilisée pour enregistrer des processus complexes et difficiles à expliquer, montrant l'interaction entre les participants du début à la fin pour l'évaluation. Ceci est particulièrement utile pour l'évaluation des interactions sociales et l'évaluation du changement de comportement et d'attitude (Corlazzoli, 2014).

Les différents types de visualisation de données non exhaustives ont aidé les populations cibles des interventions, les donateurs et les évaluateurs et évaluatrices à comprendre et à interpréter des constats (messages) qui peuvent être difficiles à identifier, comme les programmes de la consolidation de la paix (Corlazzoli, 2014). Par exemple, le Programme des Nations Unies pour le Développement (PNUD) en Chypre a créé *Mahallae*, une feuille de route d'action civique avec des histoires et des informations promouvant la paix et l'engagement civique présentées à l'aide de visualisations de données. Les informations pour ces visualisations sont soumises par les membres de la communauté, puis analysées et partagées à l'aide de visualisations faciles à comprendre (Corlazzoli, 2014). Également, *AidData* de *Development Gateway* est une source de données en ligne spécialement conçue pour les acteurs du développement international. Les données sont présentées sous la forme d'un tableau de bord en ligne et offrent aux parties prenantes, aux donateurs et aux acteurs du développement la possibilité de passer en revue les informations pertinentes pour leur intervention (Corlazzoli, 2014).

## Limites et enjeux des TIC

Les TIC présentent un énorme potentiel pour compléter et dynamiser les approches actuelles de l'évaluation, mais un certain nombre de défis existent (Kaboré, 2021). On note des défis logistiques, méthodologiques et de confidentialité des données.

- **Au niveau logistique**
  - La faible capacité institutionnelle et la résistance au changement, qui sont des défis communs aux organisations qui n'ont pas de budget nécessaire pour former et intégrer pleinement les TIC dans leurs activités.
  - Les TIC peuvent avoir des coûts initiaux élevés, ce qui signifie que les économies ne sont pas immédiatement apparentes ou que les fonds ne sont pas disponibles pour d'autres types de collecte de données.
  - Le développement et la mise en œuvre d'un outil ou d'un système de collecte de données numériques entièrement nouveau nécessitent souvent un investissement initial important ou des coûts de maintenance permanents élevés (Bamberger *et al.*, 2016; Yamamoto, 2014).

- **Au niveau méthodologique**

- La collecte de données à distance par téléphone signifie que de nombreux indices contextuels observés lors d'un entretien en face à face seront absents (Kaboré, 2021).
- Des techniques telles que le financement participatif (*crowdsourcing*) permettent de recueillir des informations et des opinions de milliers de personnes, mais l'échantillon peut être biaisé.
- Alors que de nombreux évaluateurs et évaluatrices prétendent utiliser des méthodes mixtes, dans la plupart des cas, ils n'utilisent pas une approche intégrée systématique de méthodes mixtes, mais utilisent plutôt des méthodes quantitatives et qualitatives indépendamment les unes des autres avec peu de recours systématiques à la triangulation. Par conséquent, on apprécie encore relativement peu la manière dont les TIC pourraient être utilisées pour renforcer la rigueur méthodologique (Bamberger *et al.*, 2016).

- **Au niveau de la confidentialité**

- Le stockage des données numériques peut entraîner des risques accrus pour la vie privée et la sécurité des personnes participant à l'évaluation (Kaboré, 2021).
- Bien que de nombreux indices contextuels puissent théoriquement être saisis sur les téléphones intelligents par le biais d'images audio ou visuelles, il est difficile d'expliquer la collecte de données numériques à ceux qui n'ont pas fait l'expérience des TIC, ce qui soulève des doutes quant au consentement véritablement éclairé (Bamberger *et al.*, 2016).

En définitive, l'utilisation des TIC, perçue comme un avantage, présente un certain nombre de défis. Il s'agit à la fois de problèmes techniques (mauvaise connectivité) et de problèmes « centrés sur l'humain » (considérations éthiques), qui continuent d'entraver l'utilisation généralisée des TIC (Meador, 2020).

## Conclusion

Les TIC sont devenues incontournables en évaluation. Elles présentent un énorme potentiel pour dynamiser les approches actuelles d'évaluation. Elles aident à résoudre des contraintes de temps, de ressources, de qualité des données et de biais d'échantillonnage (Kaboré, 2021). Également, les TIC permettent de mieux atteindre les populations visées dans la mise en œuvre des interventions, ce qui limiterait les problèmes d'exclusion sociale et d'équité (Bamberger *et al.*, 2016). Plusieurs limites ont été présentées, mais plutôt que de traiter les TIC comme une solution unique, les évaluateurs et évaluatrices devraient les considérer comme « une partie de la solution à certains défis » (Corlazzoli, 2014).

## Pour aller plus loin :

1. Corlazzoli, V. (2014). *ICTs for Monitoring & Evaluation of Peacebuilding Programs*. DFID : Department for International Development. <https://www.sfcg.org/wp-content/uploads/2014/05/CCVRI-SSP-ICT-and-ME-Final.pdf>
2. Meador, K. (2020). *Information and communications technologies (ICT) in program evaluation practice : Exploring new pathways to evaluation use* [thèse de doctorat, Université d'Ottawa]. [https://ruor.uottawa.ca/bitstream/10393/40556/1/Meador\\_Kateryna\\_2020\\_Thesis.pdf](https://ruor.uottawa.ca/bitstream/10393/40556/1/Meador_Kateryna_2020_Thesis.pdf)

Pour plus d'information sur les TIC en évaluation, ces documents, en anglais, présentent plusieurs outils ainsi que des liens et des ressources.

## Références bibliographiques

- Azzam, T. et Jacobson, M. R. (2015). Reflections on the Future of Research. *New Directions for Evaluation*, 148(2015), 103-116. <https://onlinelibrary.wiley.com/toc/1534875x/2015/2015/148>
- Azzam, T. et Robinson, D. (2013). GIS in evaluation : Using the power of geographic information systems to represent evaluation data. *American Journal of Evaluation*, 34(2), 207-24. <https://journals.sagepub.com/doi/abs/10.1177/1098214012461710?journalCode=ajec>
- Bamberger, M., Raftree, L. et Olazabal, V. (2016). The role of new information and communication technologies in equity-focused evaluation : opportunities and challenges. *Evaluation*, 22(2), 228-244. <https://journals.sagepub.com/doi/abs/10.1177/1356389016638598>
- Banque mondiale. (2012). *Les technologies de l'information et des communications sont en train de révolutionner le développement en Afrique*. <https://www.banquemondiale.org/fr/news/press-release/2012/12/10/information-communication-technology-revolutionizing-development-africa>
- Burt, M. (2013). The “poverty spotlight” approach to eliminating multidimensional poverty : Business, civil society, and government working together in Paraguay. *Innovations*, 8(1-2), 47-67. <https://direct.mit.edu/itgg/article/8/1-2/47/9742/The-Poverty-Stoplight-Approach-to-Eliminating>
- Corlazzoli, V. (dir.) (2014). *ICTs for Monitoring & Evaluation of Peacebuilding Programs*. DFID : Department for International Development, UK Aid for British People. <https://www.sfcg.org/wp-content/uploads/2014/05/CCVRI-SSP-ICT-and-ME-Final.pdf>
- De Miranda, C. (2019). *ICT4D : El uso de las tecnologías digitales para la promoción del desarrollo social*. Instituto del Futuro. [https://idf.pe/detail\\_blog.php?lang=es&id=25](https://idf.pe/detail_blog.php?lang=es&id=25)
- Evergreen, S. et Metzner, C. (2013). Design principles for data visualization in evaluation. *New Directions for Evaluation*, 140, 5-20. <https://doi.org/10.1002/ev.20071>
- Gobble, M. A. (2013). Big Data : The next big thing in innovation. *Research-Technology Management*, 56(1). <https://www.tandfonline.com/doi/abs/10.5437/08956308X5601005>
- Gossart, C. (2010). «TIC et développement durable : Les conditions du succès», de Christelle Hue, Isabelle Nicolaï et Sylvie Faucheux (2010), De Boeck. *Terminal*, 106-107. <https://journals.openedition.org/terminal/1827#quotation>



- Griffiths, E. (2012). Geographic Information Systems (GIS) and Spatial Analysis. Dans M. Williams et W. P. W. Vogt (dir.), *The SAGE Handbook of Innovation in Social Research Methods* (p. 442-464). SAGE Publications.
- Hites, L. S., Fifolt, M., Beck, H., Su, W., Kerbawy, S., Wakelee, J. et Nassel, A. (2013). A Geospatial Mixed Methods Approach to Assessing Campus Safety. *Evaluation Review*, 37(5), 347-369. <https://journals.sagepub.com/doi/abs/10.1177/0193841X13509815?journalCode=erxb>
- Humanitarian Practice Network et Scriven, K. (2016). Humanitarian Innovation. *Humanitarian Exchange Magazine*, 66, 1-47. <https://odihpn.org/magazine/humanitarian-innovation/>
- IFAD. (2017, 6-7 juin). *Information and Communication Technologies for Evaluation (ICT4Eval) international conference : Using innovative approaches to development evaluation International conference* [communication orale]. Independent Office of Evaluation, IFAD, Rome, Italie. <https://www.ifad.org/en/web/ioe/event/asset/39824138>
- Kaboré, T. P. (2021). *Pratique de l'évaluation : Quel avenir à l'ère des technologies de l'information et de communication pour le développement?* (Note de synthèse n°16). Réseau Francophone de l'Évaluation. <https://drive.google.com/file/d/1J6oHE5WRrQsPZg9MNtZMOfoTpkAaLj3S/view>
- Lefebvre-Nare, F., Lemire, S. et Petersson, G. J. (2017). What is Big Data? Dans G. J. Petersson et J. D. Breul (dir.), *Cyber society, big data, and evaluation* (p. 19-34). NJ : Transaction Publishers.
- Mazzeo Rinaldi, F., Giuffrida, G. et Negrete, T. (2017). Real-time monitoring and evaluation – emerging news as predictive process using Big Data-based approach. Dans G. J. Petersson et J. D. Breul (dir.), *Cyber Society, big data, and evaluation* (p. 191-213). NJ : Transaction Publishers.
- Meador, K. (2020). *Information and communications technologies (ICT) in program evaluation practice : exploring new pathways to evaluation use* [thèse de doctorat, Université d'Ottawa]. [https://ruor.uottawa.ca/bitstream/10393/40556/1/Meador\\_Kateryna\\_2020\\_Thesis.pdf](https://ruor.uottawa.ca/bitstream/10393/40556/1/Meador_Kateryna_2020_Thesis.pdf)
- Nediger, M. (2019, 19 juin). *Qu'est-ce qu'une infographie? Exemples d'infographies, modèles et astuces de design*. Venngage. <https://fr.venngage.com/blog/qu-est-ce-qu-une-infographie/>
- Petersson, G. J., Leeuw, F., Breul, J. et Leeuw, H. B. (2017a). Cyber society, Big Data, and evaluation : an Introduction. Dans G. J. Petersson et J. Breul (dir.), *Cyber Society, Big Data, and Evaluation* (p. 1-17). NJ : Transaction Publishers.
- Raftree, L. (2014, 14 mars). *Wait... What? Discussions on digital ethics, privacy and power*. ICTs and M&E at the African Evaluators' Conference (Part 1). <https://lindaraftree.com/2014/03/14/icts-and-me-at-the-african-evaluators-conference-part-1/>
- Raftree, L. et Bamberger, M. (2014). *Emerging Opportunities : Monitoring and Evaluation in a Tech-Enable World*. The Rockefeller Foundation : evaluation office. <https://www.rockefellerfoundation.org/wp-content/uploads/Monitoring-and-Evaluation-in-a-Tech-Enabled-World.pdf>
- Social Media*. (2020). Merriam-Webster. <https://www.merriam-webster.com/dictionary/social%20media>
- Yamamoto, G. (2015). *Technology – and evaluation of market-based solutions : A multi-country impact study in 3 countries*. Opportunity International and MasterCard Foundation. Communication présentée à la conférence annuelle de l'American Evaluation Association, Denver.



# Partie III

## Phase empirique et analytique



# L'implication des parties prenantes en évaluation



*Kris Essosola Eale  
Mary Pat Selvaggio*

## Introduction

Le chapitre « Les parties prenantes en évaluation : généralités » donne quelques clés pour identifier les parties prenantes de l'évaluation. Il devient alors important de déterminer comment les impliquer et à quel stade de l'évaluation le faire. Ce chapitre aborde les défis associés à l'implication des parties prenantes dans la démarche évaluative en fonction des approches connues, la combinaison des atouts de l'approche participative et de l'évaluation conventionnelle, ainsi que des enseignements tirés et des stratégies utilisées provenant de nos pratiques ainsi que de celles de nos collègues.

## Défis associés à l'implication des parties prenantes dans la démarche évaluative

Tout évaluateur ou évaluatrice doit comprendre que les parties prenantes peuvent à la fois aider ou entraver une évaluation à tout stade de la démarche évaluative. Les évaluateurs et évaluatrices sur le terrain peuvent témoigner des difficultés survenues à la suite de l'inaccessibilité des parties prenantes, de leur refus ou d'autres limitations de participer, de leur manque d'intérêt ou même, cas extrême, de leur rejet des résultats. D'où la nécessité de comprendre à quel stade de la démarche évaluative chaque partie prenante est requise et l'importance de prendre des précautions pour les impliquer au bon moment. Haddaway *et al.* (2017) ont noté que l'engagement des parties prenantes a été associé à un certain nombre de défis qui rendent la mise en œuvre de celui-ci problématique, notamment : une demande accrue de ressources et de temps, le potentiel de marginaliser ou de favoriser certains groupes de parties prenantes, une représentation biaisée des véritables groupes de parties prenantes et le risque d'engagement symbolique (Haddaway *et al.*, 2017).

Cependant, d'après notre expérience, un défi plus important pour faire participer les parties prenantes à l'évaluation réside dans la sollicitation initiale des évaluations. Souvent, les organisations qui commandent des évaluations (bailleurs de fonds ou partenaires de mise en œuvre) ne traitent pas de l'engagement des parties prenantes dans les termes de référence (TdR)/ demande de propositions (DP).

La sélection des évaluateurs, des évaluatrices ou de l'équipe d'évaluation dépend également de la même organisation qui commande l'évaluation, les critères de sélection étant souvent basés sur l'alignement de la proposition soumise avec le TdR/DP, en lien notamment avec la solidité de la méthodologie suggérée et de l'approche proposée (qui reflètent souvent leur demande de propositions), de l'expérience technique de l'équipe d'évaluation recommandée, de la performance antérieure de l'entreprise, ainsi que de l'alignement du budget proposé sur le budget plafond de la demande de propositions. Tout ce qui précède laisse peu de marge de manœuvre aux évaluateurs et évaluatrices, car pour remporter la soumission, ils doivent concevoir leurs propositions en fonction de la demande de propositions.

De plus, après avoir remporté l'offre, les évaluateurs et évaluatrices doivent travailler en étroite collaboration avec l'organisation qui commande l'évaluation pour avoir accès aux données et aux participants. Parfois, le personnel de l'organisation qui commande l'évaluation peut appréhender l'indépendance des évaluateurs et évaluatrices ainsi que l'orientation de la responsabilité de l'évaluation. Ceci crée souvent une relation de travail tendue et inconfortable, qui peut conduire à un manque de coopération et à une perception erronée des évaluatrices et évaluateurs en tant qu'auditeurs/enquêteurs qui ont été embauchés pour exposer les faiblesses de l'organisation. Dans ces cas, les évaluateurs et évaluatrices sont souvent sous pression et essaient de gagner la coopération du personnel en s'adaptant à ses limites, ce qui implique quelquefois de faire preuve de trop de flexibilité, et souvent de compromettre leurs méthodologies et leurs approches. Ainsi, les évaluateurs et évaluatrices peuvent accepter de n'accéder qu'aux parties prenantes dictées par les organisations qui commandent ces évaluations.

À partir de ce qui précède, le processus de sollicitation et de sélection des évaluations peut préparer le terrain pour compromettre l'indépendance des évaluateurs et évaluatrices ainsi que leur capacité à plaider en faveur d'un processus plus participatif qui identifie et engage de manière indépendante les principales parties prenantes à chaque étape de la démarche évaluative. En fait, lorsque des mandats ou des appels d'offres sont rédigés sans la consultation de toutes les parties prenantes concernées et que la conduite de l'évaluation est menée sous la direction de la même organisation qui a commandé l'évaluation, l'indépendance de l'évaluateur ou de l'évaluatrice est limitée.

Les Normes de qualité pour l'évaluation du développement de l'Organisation de coopération et de développement économiques (OCDE) soulignent l'importance de l'indépendance des évaluateurs et évaluatrices vis-à-vis de toutes les parties prenantes, y compris les partenaires d'exécution et les agences de financement. L'OCDE insiste sur le fait que tous les conflits d'intérêts possibles sont traités ouvertement et honnêtement, et que l'équipe d'évaluation doit être en mesure de travailler librement et sans interférence et être assurée de la coopération et de l'accès à toutes les informations pertinentes (OCDE, 2010, p. 11).

Il est encourageant de constater que les règles de certaines agences de financement, telles que la politique d'évaluation et les lignes directrices pour les évaluations du ministère néerlandais des Affaires étrangères, exigent que le projet du TdR soit discuté à l'avance avec les parties prenantes directes, avec des experts externes (de préférence dans la forme d'un groupe de référence) et avec les pairs examinateurs internes (Dutch Ministry of Foreign Affairs, 2009, p. 31).

## Limites des approches théoriques dans l'engagement des parties prenantes dans la démarche évaluative

L'OCDE fixe deux objectifs clés à une évaluation (OCDE, 2010, p. 8) :

- **Établir la responsabilité**—pour rendre compte des résultats obtenus avec les ressources allouées pour mettre en œuvre l'intervention;
- **Apprendre**—apprendre de l'expérience en comprenant si l'intervention mise en œuvre a fonctionné ou non et les raisons de son échec ou de son succès. Contribuer à améliorer une politique, une procédure ou une technique de développement et informer la poursuite ou l'arrêt de l'intervention.

Dans la coopération internationale, il fut un temps où les évaluations avaient tendance à être gérées avec une perspective extérieure, donnant souvent peu de reconnaissance aux attentes locales et au potentiel de contributions des parties prenantes (Mcloughlin et Walton, 2012). Dans ces évaluations « conventionnelles », la direction de l'évaluation est plutôt assumée par les bailleurs de fonds et les gestionnaires des interventions. Les évaluatrices et évaluateurs professionnels déterminent les indicateurs de progrès de l'intervention en jouant un rôle d'expert ou de leader. La nécessité d'un jugement indépendant, d'une information spécialisée que seuls des experts peuvent fournir ou d'indicateurs standardisés (par exemple, reconnus au niveau mondial ou dans les sciences) sont autant de critères/bénéfices qui orientent ce choix (Zukoski et Luluquisen, 2002). Cependant, les parties prenantes sont alors des objets d'évaluation plutôt que des participants clés. Les bénéficiaires, les organisations locales et les gouvernements sont laissés sans rôle substantiel (Mcloughlin et Walton, 2012, p.12). Même si les coûts associés aux évaluations conventionnelles sont alors relativement simples à établir, puisqu'il s'agit des honoraires de consultants et d'experts, la possibilité de perdre des informations critiques que seules les parties prenantes peuvent fournir demeure (par exemple lorsqu'elles n'auraient même pas été impliquées à l'étape de la collecte de données). Dans ce contexte, l'évaluation participative est louée, car elle permet d'impliquer les principales parties prenantes en tant que participants actifs, et offre de nouvelles méthodes plus inclusives d'évaluation et d'apprentissage, reflétant les perspectives et les aspirations des personnes les plus directement touchées. Avec l'accent croissant mis sur les approches participatives du développement, il a été reconnu que l'évaluation devrait également être participative (Dillon, 2020).

Dans la famille des approches participatives, la direction de l'évaluation est plutôt de type collégial et implique jusqu'au personnel des interventions ainsi que les bénéficiaires. Ces mêmes parties prenantes pourront déterminer les indicateurs de progrès de l'intervention en collaboration avec l'équipe d'évaluation qui agit alors comme un *coach*, un facilitateur, un négociateur, un « ami critique ». Ces approches sont particulièrement utiles dans les situations où des difficultés sur la mise en œuvre de l'intervention sont apparues; lorsque des questions d'effets touchent particulièrement les populations cibles (les bénéficiaires), mais également celles qui ont été exposées à l'intervention ou pour répondre à un besoin spécifique d'information d'une partie prenante. Parmi les bénéfices, on peut alors citer la valorisation du savoir local, l'augmentation de la fiabilité des informations et des constats, ainsi que le renforcement des capacités des acteurs impliqués en vue de leur autonomisation (Zukoski et Luluquisen, 2002).

Pourtant, l'évaluation participative a ses défis. Cullen note qu'il y a moins de consensus sur ce que l'on entend par évaluation participative que sur le sujet des parties prenantes (Cullen, 2009, p. 25). En effet, l'éventail des méthodes classées comme participatives varie considérablement (voir l'extrait du tableau 1 de la classification de celles-ci par Cullen). Pour certains, ce sont celles qui impliquent tout type de consultation avec les parties prenantes. Pour d'autres, une évaluation n'est véritablement participative que lorsque les principales parties prenantes sont activement impliquées à toutes les étapes de l'évaluation. Cousins note que la signification, la nature et la forme mal définies de l'évaluation participative conduisent à des confusions autour des conditions dans lesquelles elle est la plus appropriée et des conséquences auxquelles elle peut mener (Cousins, 2003, p. 245). Guijt et Gaventa affirment fermement que l'évaluation participative ne consiste pas seulement à utiliser des techniques participatives dans un cadre conventionnel de suivi et d'évaluation. Il s'agit de repenser radicalement qui initie et entreprend la démarche évaluative, et qui apprend ou profite des résultats (Guijt et Gaventa 1998, p. 2). Cette définition radicale a contribué à établir l'approche d'évaluation participative comme une approche d'évaluation autonome avec sa propre méthodologie systématique et ses propres exigences, la rendant presque inaccessible et difficile à mettre en œuvre. En conséquence, les évaluateurs et évaluatrices n'ont pas utilisé ces techniques, pourtant utiles, dans la démarche évaluative conventionnelle. D'après notre expérience en évaluation, l'approche d'évaluation participative radicale n'est pas utilisée par les évaluateurs et évaluatrices dans la pratique, car elle nécessite beaucoup de temps, d'énergie et d'engagement de la part de toutes les parties prenantes, ainsi que la coordination de nombreux acteurs, le développement de la formation et des compétences et le soutien pour l'engagement des parties prenantes. Cela ne correspond pas au temps généralement court alloué pour mener l'évaluation. L'évaluation participative peut également conduire à des conflits potentiels entre les parties prenantes qui peuvent allonger encore davantage le processus, comme le notent Zukoski et Luluquisen : « Les évaluations participatives nécessitent une planification de la résolution des conflits entre les individus impliqués. Les conflits peuvent survenir en raison de différences culturelles, linguistiques, de classe et autres qui existent entre et au sein des groupes. Ces conflits peuvent entraver le travail d'équipe requis pour une évaluation participative » (Zukoski et Luluquisen 2002, p. 6).

Étant donné que la plupart des évaluations sont limitées en temps et en budget, ce qui nécessite un traitement très rapide des informations, notre conseil pratique est que les évaluateurs et évaluatrices ne se perdent pas dans les différents cadres théoriques et les approches de l'engagement des parties prenantes, mais qu'ils les utilisent pour mieux adapter leurs méthodologies dans une perspective d'inclusion des parties prenantes en fonction de leur contexte d'évaluation. En d'autres termes, les évaluateurs et évaluatrices doivent déterminer la nécessité et le degré approprié d'implication de chaque partie prenante clé en fonction de l'importance de leur contribution à chaque étape de l'évaluation et de l'utilisation des résultats de l'évaluation par celle-ci.

**Tableau 1 — Classification des approches participatives. Extrait reproduit de Cullen, 2009, p. 48, traduit librement par les auteurs avec la permission de Cullen.**

Nom des approches ou méthodes	Auteur(s) principal(aux)	Objectif de la participation	Contrôle de la prise de décision	Degré de participation
<b>Évaluation participative pratique (EPI)</b>	Cousins et Earl; Ayers.	Pratique : aide à la prise de décision de l'intervention et à la résolution de problèmes; utilisation de l'évaluation.	Équilibré : évaluateur ou évaluatrice et participants en partenariat.	Large : participation à toutes les phases de l'évaluation.
<b>Évaluation participative transformatrice (TPE)</b>	Tandon et Fernandes; Fals-Borda; Gaventa.	Politique : autonomisation, émancipation, justice sociale.	Équilibré : partenariat, mais contrôle ultime de la prise de décision par les participants.	Large : participation à toutes les phases de l'évaluation.
<b>Évaluation basée sur les parties prenantes</b>	Bryk; Mark et Shotland.	Pratique : utilisation de l'évaluation; un certain accent mis sur les aspects politiques de l'évaluation.	Évaluateur ou évaluatrice : coordinateur des activités et des aspects techniques de l'évaluation.	Limité : parties prenantes consultées lors des phases de planification et d'interprétation.
<b>Évaluation en milieu scolaire</b>	Nevo; Alvik.	Pratique : aide à la prise de décision de l'intervention et à la résolution de problèmes.	Équilibré : l'évaluateur ou l'évaluatrice forme le personnel de l'école qui fait sa propre enquête.	Large : participation à toutes les phases de l'évaluation.
<b>Évaluation démocratique</b>	MacDonald; McTaggart.	Politique : utilisation légitime de l'évaluation dans une société pluraliste.	Équilibré : l'évaluateur ou l'évaluatrice et les participants travaillent en partenariat.	Modéré : les parties prenantes contrôlent l'interprétation et les rapports.
<b>Évaluation évolutive</b>	Patton	Pratique : amélioration de l'intervention; utilisation de l'évaluation.	Équilibré : l'évaluateur ou l'évaluatrice et les participants travaillent en partenariat.	Important : implication et participation continue.
<b>Évaluation de l'autonomisation</b>	Fetterman	Politique : autonomisation, illumination, autodétermination.	Participants : contrôle presque complet, facilité par l'évaluateur ou l'évaluatrice.	Large : participation à toutes les phases de l'évaluation.
<b>Évaluation axée sur l'utilisation</b>	Patton	Pratique : utilisation de l'évaluation.	Équilibré : évaluateur ou évaluatrice et participants en partenariat.	Large : participation à toutes les phases de l'évaluation.



Nom des approches ou méthodes	Auteur(s) principal(aux)	Objectif de la participation	Contrôle de la prise de décision	Degré de participation
<b>Évaluation répondante</b>	Stake	Politique/ philosophique : promouvoir l'équité et la justice, sensible aux préoccupations des parties prenantes.	Évaluateur ou évaluatrice : maintient le contrôle et l'autorité.	Large : participation à toutes les phases de l'évaluation, en particulier pour identifier les problèmes de l'intervention.
<b>Évaluation collaborative</b>	Rodriguez-Campos	Pratique : soutien à la prise de décision de l'intervention, appropriation partagée, qualité accrue.	Équilibré : l'évaluateur ou l'évaluatrice et les membres de la collaboration travaillent en partenariat.	Large : les membres de la collaboration travaillent ensemble à toutes les phases de l'évaluation.
<b>Évaluation rurale rapide</b>	Chambers; Dart.	Pratique/ philosophique : moyen rentable de collecter des informations auprès des populations locales.	Évaluateur ou évaluatrice : bien que les bénéficiaires participent à la collecte des données, l'évaluateur ou l'évaluatrice garde le contrôle.	Modéré : participation principalement limitée à la collecte de données.
<b>Évaluation rurale participative</b>	Chambers; Townsley.	Politique/ philosophique : mouvement vers les préoccupations des « initiés » par rapport aux « étrangers » dans le processus de développement.	Praticien : les rôles de l'enquêté et de l'enquêteur sont inversés.	Large : les parties prenantes participent largement à tous les aspects.
<b>Évaluation des bénéficiaires</b>	World Bank	Pratique/ philosophique : en écoutant les personnes marginalisées, les responsables de l'intervention peuvent prendre de meilleures décisions.	Évaluateur ou évaluatrice	Modéré : cependant, l'évaluateur ou l'évaluatrice passe beaucoup de temps à connaître les problèmes et les préoccupations des parties prenantes.

## **Comment impliquer les parties prenantes dans l'évaluation en combinant les atouts de l'approche participative et de l'évaluation conventionnelle ?**

Alors que de nombreux évaluateurs et évaluatrices se positionnent pour ou contre l'approche participative radicale ou l'approche conventionnelle, nous adoptons normalement une approche modérée qui combine les forces des deux approches. Cela permet de raccourcir le long processus d'évaluation participative radicale, tout en garantissant un engagement significatif des parties prenantes dans un délai raisonnable et un budget limité.

L'évaluateur ou l'évaluatrice joue ici le rôle de conducteur, d'arbitre et de médiateur dans la démarche évaluative. La responsabilité de l'évaluateur ou de l'évaluatrice n'est pas seulement d'utiliser son expertise pour guider la méthodologie et l'approche de l'évaluation, mais surtout d'identifier, de préconiser et d'assurer une participation équitable de toutes les parties prenantes clés tout au long du processus d'évaluation, lorsque leur contribution est nécessaire et requise.

Dans le tableau ci-dessous, nous proposons des pistes d'implication des parties prenantes à partir des neuf étapes proposées dans le guide du gestionnaire de l'évaluation de BetterEvaluation (s. d.-a, s. d.-b) et adapté dans le chapitre « La gestion d'une évaluation » de cet ouvrage. Cela dit, il incombe aux évaluateurs et évaluatrices d'identifier toutes les parties prenantes et, dans un consensus, de convenir du degré d'implication que chaque catégorie de parties prenantes devrait avoir en fonction de leur contribution requise (BetterEvaluation, s. d.-c, s. d.-d).

**Tableau 2** — Étapes du processus de commande d'évaluation et parties prenantes à impliquer. À partir de l'expérience des auteurs et du guide du gestionnaire d'évaluation de BetterEvaluation (s. d.-a, s. d.-b, s. d.-c, s. d.-d).

Étapes de la gestion d'une évaluation (chapitre 5 : <i>La gestion d'une évaluation</i> )	Les étapes de BetterEvaluation	Sous-étapes proposées par BetterEvaluation	Les parties prenantes à impliquer
<b>Formes, finalités et faisabilité de l'évaluation</b>	1. Déterminer de quelle façon les décisions concernant l'évaluation seront prises.	<ul style="list-style-type: none"> <li>▪ Identifier quelles personnes seront impliquées dans les décisions et quels seront leurs rôles;</li> <li>▪ Préciser les responsabilités du responsable de l'évaluation et des évaluateurs et évaluatrices;</li> <li>▪ Traiter des problèmes particuliers de gestion des évaluations liés aux interventions conjointes, y compris les partenariats avec les donateurs.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Évaluatrices et évaluateurs indépendants (arbitre) : servent de médiateur pour s'assurer que toutes les voix des parties prenantes sont intégrées où cela est nécessaire;</li> <li>▪ Partenaires de mise en œuvre;</li> <li>▪ Organisation de financement;</li> <li>▪ Représentants des bénéficiaires.</li> </ul>
	2. Définir la portée de l'évaluation.	<ul style="list-style-type: none"> <li>▪ Clarifier ce qui sera évalué;</li> <li>▪ Décrire la théorie du changement;</li> <li>▪ Identifier qui sont les principaux utilisateurs prévus de l'évaluation et de quelle façon ils l'utiliseront;</li> <li>▪ Élaborer des questions d'évaluation clés inclusives;</li> <li>▪ Décider du bon moment pour l'évaluation;</li> <li>▪ Décider si l'évaluation sera faite par une équipe externe, une équipe interne ou un hybride des deux;</li> <li>▪ Déterminer les qualités de l'évaluateur ou de l'évaluatrice;</li> <li>▪ Identifier les ressources disponibles pour l'évaluation et ce qui sera nécessaire (p. ex. nombre de jours multiplié par les honoraires d'une personne à l'externe) pour réaliser le mandat.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Évaluatrices et évaluateurs indépendants (arbitres) : servent de médiateur pour s'assurer que toutes les voix des parties prenantes sont intégrées où cela est nécessaire;</li> <li>▪ Partenaires de mise en œuvre;</li> <li>▪ Organisation de financement;</li> <li>▪ Représentants des bénéficiaires.</li> </ul>

Étapes de la gestion d'une évaluation (chapitre 5 : <i>La gestion d'une évaluation</i> )	Les étapes de BetterEvaluation	Sous-étapes proposées par BetterEvaluation	Les parties prenantes à impliquer
<b>Étude préparatoire et termes de références</b>	3. Élaborer le mandat.	<ul style="list-style-type: none"> <li>▪ Spécifier le contexte de l'évaluation;</li> <li>▪ Indiquer les buts/objectifs /justification de l'évaluation;</li> <li>▪ Préciser les utilisateurs et les utilisations prévues de l'évaluation;</li> <li>▪ Spécifier les questions d'évaluation clés;</li> <li>▪ Souligner les principes et l'approche qui guideront l'évaluation;</li> <li>▪ Spécifier les méthodologies;</li> <li>▪ Mentionner les rôles et responsabilités des différents acteurs;</li> <li>▪ Spécifier les exigences en matière de rapports;</li> <li>▪ Énoncer la chronologie et les jalons de l'évaluation;</li> <li>▪ Spécifier toute autre exigence spécifique.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Évaluatrices et évaluateurs indépendants (arbitres) : servent de médiateur pour s'assurer que toutes les voix des parties prenantes sont intégrées où cela est nécessaire;</li> <li>▪ Partenaires de mise en œuvre;</li> <li>▪ Organisation de financement;</li> <li>▪ Représentants des bénéficiaires.</li> </ul>
<b>Constitution de l'équipe d'évaluation</b>	4. Engager l'équipe d'évaluation.	<ul style="list-style-type: none"> <li>▪ Annoncer les termes de référence de l'évaluation (TdR) / la demande de propositions (DP);</li> <li>▪ Sélectionner un évaluateur, une évaluatrice ou une équipe d'évaluation;</li> <li>▪ Préparer le contrat;</li> <li>▪ Orienter l'évaluateur, l'évaluatrice ou l'équipe d'évaluation.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Évaluatrices et évaluateurs indépendants (arbitres) : devraient publier le mandat, sélectionner une équipe d'évaluatrices et d'évaluateurs indépendants et guider l'orientation de l'équipe d'évaluation;</li> <li>▪ Partenaire de mise en œuvre/organisme de financement : préparer le contrat et apporter sa contribution à l'orientation.</li> </ul>

Étapes de la gestion d'une évaluation (chapitre 5 : <i>La gestion d'une évaluation</i> )	Les étapes de BetterEvaluation	Sous-étapes proposées par BetterEvaluation	Les parties prenantes à impliquer
<b>Lancement de l'évaluation</b>	5. Gérer la conception de l'évaluation (méthodologie). 6. Gérer l'élaboration du plan de travail d'évaluation, y compris la logistique.	<ul style="list-style-type: none"> <li>▪ Élaboration du rapport de démarrage :                             <ul style="list-style-type: none"> <li>- Mettre en évidence la méthodologie et les approches d'évaluation finale, y compris la conception des outils, l'échantillonnage des sites, etc.;</li> <li>- Tenir compte des implications des ressources disponibles et des contraintes spécifiques. Cela comprend le plan de travail, la logistique et le niveau d'effort de chaque membre de l'équipe.</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>▪ Équipe d'évaluation sélectionnée : élaboration de la conception de l'évaluation;</li> <li>▪ Partenaire de mise en œuvre / organisme de financement : fournir des commentaires et des contributions (examen).</li> </ul>
<b>Mise en œuvre de l'évaluation</b>	7. Gérer la mise en œuvre de l'évaluation.	<ul style="list-style-type: none"> <li>▪ Formation de l'équipe des agents de terrain;</li> <li>▪ Travail de terrain : collecte de données;</li> <li>▪ Nettoyage des données;</li> <li>▪ Analyse des données (production d'analyses).</li> </ul>	<ul style="list-style-type: none"> <li>▪ Équipe d'évaluation sélectionnée : recueillir les contributions des parties prenantes.</li> <li>▪ Partenaires de mise en œuvre / organisation de financement : fournir des contributions à l'équipe d'évaluation;</li> <li>▪ Bénéficiaires : fournir des contributions à l'équipe d'évaluation.</li> </ul>
<b>Élaboration du rapport</b>	8. Guider la production de rapports de qualité.		<ul style="list-style-type: none"> <li>▪ Équipe d'évaluation sélectionnée : écriture du rapport et présentation des résultats;</li> </ul>
<b>Soumission et discussion du rapport</b> <b>Plan d'action de suivi des recommandations</b> <b>Dissémination des leçons apprises</b>	9. Diffuser les rapports et appuyer l'utilisation de l'évaluation.		<ul style="list-style-type: none"> <li>▪ Partenaire de mise en œuvre / organisme de financement : fournir des commentaires et des contributions (examen);</li> <li>▪ Représentants des bénéficiaires : fournissent des commentaires et des contributions.</li> </ul>

## **Formes, finalités, faisabilité de l'évaluation, études préparatoires et termes de référence : une phase de consultation prémandat/TdR**

L'indépendance de l'évaluation commence à ce stade. Nous sommes d'accord avec la politique d'évaluation et les lignes directrices pour les évaluations du ministère néerlandais des Affaires étrangères selon lesquelles la création du mandat (appel d'offres d'évaluation) devrait impliquer au maximum les parties prenantes concernées (Dutch Ministry of Foreign Affairs, 2009). Cela garantira que les parties prenantes s'approprient l'évaluation, que l'évaluation est pertinente, et assurera que l'évaluation a une utilité maximale pour toutes les parties prenantes, et pas seulement l'agence qui commande l'évaluation. En outre, le TdR/DP doit garantir que la crédibilité et la fiabilité de l'évaluation sont renforcées en garantissant que les propositions décrivent comment les parties prenantes seront engagées et quelle méthodologie sera utilisée. Une fois que le TdR/DP est validé par toutes les parties prenantes concernées, il peut être communiqué aux soumissionnaires indépendants potentiels.

### **Constitution de l'équipe d'évaluation**

Dans l'idéal, l'examen des propositions et la sélection de la proposition gagnante devraient se faire par une organisation indépendante, indépendamment des organismes de financement et de mise en œuvre. Cela renforce encore l'indépendance de l'évaluation. L'indépendance de l'équipe d'évaluation ne signifie pas que les évaluateurs et évaluatrices doivent travailler de manière isolée, mais doivent être exempts de toute influence qui les exposerait à des préjugés ou à une flexibilité excessive envers une partie prenante particulière.

### **Lancement de l'évaluation**

La phase de lancement permettra aux évaluateurs et évaluatrices de finaliser la méthodologie et l'approche d'évaluation, y compris l'identification des répondants qui représentent toutes les parties prenantes concernées. La phase de démarrage nécessite une collaboration étroite avec le partenaire d'exécution pour finaliser la méthodologie et les approches ainsi que la planification des activités. Si la phase de consultation pré-TdR/DP a recueilli des contributions des parties prenantes suffisantes pour permettre à l'équipe d'évaluation sélectionnée de rédiger un rapport de démarrage qui soit inclusif pour toutes les parties prenantes, cette étape nécessitera très peu ou pas de consultations de toutes les autres parties prenantes.

### **Mise en œuvre de l'évaluation**

À ce stade, les parties prenantes ont la possibilité de fournir leurs contributions par le biais d'entretiens et d'autres preuves documentées. En outre, les évaluateurs et évaluatrices peuvent choisir d'engager des représentants des parties prenantes pour servir de collecteurs de données, mais des mesures doivent être mises en place pour réduire les conflits d'intérêts. Par exemple, nous avons utilisé le personnel du gouvernement de district pour collecter des données sur l'intervention en cours d'évaluation, mais pas dans leur district – ils ont été chargés de collecter des données uniquement dans un autre district où ils n'avaient ni responsabilités ni relations. Cela a également aidé à valider les résultats de l'évaluation.

## Élaboration, soumission et discussion du rapport

L'équipe d'évaluation rédige le rapport d'évaluation après avoir analysé et intégré toutes les entrées. Les parties prenantes devraient avoir la possibilité de faire part de leurs commentaires sur les résultats de l'évaluation et ceux-ci devraient être intégrés dans le rapport final. Dans notre pratique d'évaluation, nous engageons régulièrement les parties prenantes dans un « atelier de recommandations » au cours duquel les principales conclusions de l'évaluation sont présentées et les parties prenantes prennent l'initiative de formuler des recommandations. Encore une fois, cela augmente l'appropriation et l'utilité des conclusions et des recommandations de l'évaluation.

**Tableau 3** — Démarche pour combiner les atouts de l'évaluation participative et conventionnelle, d'après les questions de Zukoski et Luluquisen (2002).

Questions	Atouts combinés des évaluations participatives et conventionnelles
<b>Qui dirige l'évaluation ?</b>	Les évaluateurs et évaluatrices jouent le rôle de médiateur, de facilitateur et de <i>coach</i> pour s'assurer que toutes les parties prenantes sont impliquées et que leurs contributions sont intégrées aux résultats de l'évaluation.
<b>Qui détermine les indicateurs de progrès de l'intervention ?</b>	Toutes les parties prenantes guidées par des évaluateurs et évaluatrices indépendants.
<b>Qui est responsable de la collecte, de l'analyse des données et de la préparation des rapports finaux ?</b>	L'évaluateur ou l'évaluatrice, avec la participation et les contributions de toutes les parties prenantes.
<b>Quel est le rôle de l'évaluateur local ?</b>	<i>Coach</i> , facilitateur, négociateur, médiateur.
<b>Quels sont les bénéfices ?</b>	<ul style="list-style-type: none"> <li>▪ Offre un jugement indépendant;</li> <li>▪ Les indicateurs standardisés permettent une comparaison avec d'autres résultats de recherche;</li> <li>▪ Inclus les connaissances (savoir) locales;</li> <li>▪ Permet la vérification des informations des principaux acteurs (fiabilité);</li> <li>▪ Renforce les connaissances, les compétences et les relations entre les bénéficiaires et les autres parties prenantes.</li> </ul>

La combinaison des atouts de l'approche participative et conventionnelle permet aux évaluateurs et évaluatrices de tirer parti des avantages d'une participation significative des parties prenantes tout en conduisant l'évaluation dans un court délai et avec un budget limité. Cela permet également aux parties prenantes de se faire entendre dans les conclusions et recommandations de l'évaluation.



## Enseignements tirés et stratégies utilisées pour favoriser la participation des principales parties prenantes dans la démarche évaluative

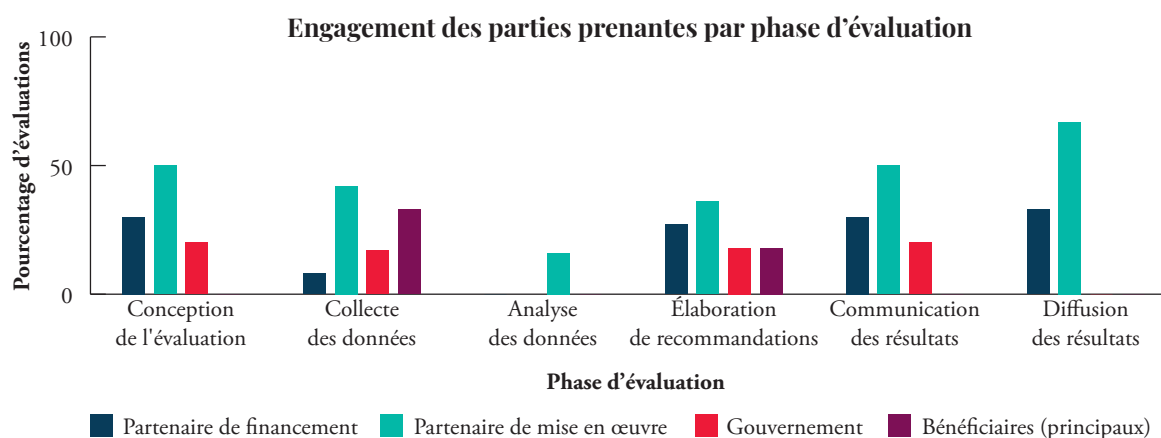
Nous avons mené une courte enquête auprès de nos collègues évaluateurs de Khulisa Management Services pour recueillir leurs points de vue sur l'implication des parties prenantes dans 12 évaluations menées par Khulisa entre 2018 et 2020. Cette enquête avait pour objectif de permettre une réflexion et de tirer des leçons pratiques sur l'implication des parties prenantes à travers l'expérience réelle des évaluateurs et évaluatrices. Nous avons posé des questions pertinentes à ce sujet pour soutenir, avec des preuves scientifiques, notre expertise partagée. Nos résultats d'enquête sont partagés ci-dessous et comparés aux phases d'évaluation de Cullen (Cullen, 2009, p. 98).

### Pendant la phase de commande des évaluations

La plupart des évaluations entreprises par Khulisa ont été commandées par des organismes de financement (75 %) ou des partenaires d'exécution (25 %). Alors que la plupart des TdR/DP (67 %) exigeaient la participation des parties prenantes au processus d'évaluation, seule la moitié des TdR (50 %) exigeait spécifiquement l'inclusion des principaux bénéficiaires.

### Implication des parties prenantes dans la démarche évaluative

Notre enquête confirme les conclusions de Cullen (2009) selon lesquelles les partenaires de mise en œuvre et les organismes de financement sont les plus impliqués dans les évaluations (à toutes les étapes du processus), les gouvernements et les principaux bénéficiaires étant moins susceptibles d'être impliqués, sauf pendant la collecte de données (moins de la moitié des évaluations) et éventuellement l'élaboration de recommandations (moins d'un quart des évaluations). Toutefois, l'engagement n'a jamais dépassé 50 % des évaluations, sauf pour la diffusion des résultats, où les partenaires d'exécution étaient les plus impliqués.



**Figure 1**

Engagement des parties prenantes par phase d'évaluation.

Source : à partir des données d'un sondage de Khulisa, 2020.

Nos collègues de Khulisa ont cité de nombreuses raisons qui empêchent spécifiquement l'implication des bénéficiaires, notamment l'incapacité à atteindre les bénéficiaires, le manque de capacité des bénéficiaires ou les contraintes budgétaires et de temps en raison des délais d'évaluation.

Lors de l'analyse des données et de la rédaction du rapport, Cullen (2009) et les évaluateurs de Khulisa ont tous affirmé que les évaluateurs et évaluatrices travaillent souvent seuls avec peu ou pas d'implication des parties prenantes à ce stade. Cependant, après analyse, lors de l'élaboration des recommandations, de la communication des résultats et de la diffusion des résultats, les parties prenantes sont susceptibles d'être consultées.

## Défis et leçons apprises

Le tableau ci-dessous résume les défis auxquels nos collègues de Khulisa sont confrontés, et démontre comment ils ont réussi à les surmonter (s'ils l'ont fait), et les leçons qu'ils ont apprises.

**Tableau 4 — Défis, stratégies d'atténuation et leçons apprises pour relever les défis d'implication des parties prenantes dans la démarche évaluative. À partir d'un sondage de Khusila, 2020.**

Quels défis à relever pour impliquer les parties prenantes dans la démarche évaluative	Comment avez-vous surmonté ces défis ?	Quelles leçons avez-vous tirées concernant l'inclusion des parties prenantes, en particulier des bénéficiaires, dans la démarche évaluative ?
<b>Taux de réponse bas des bénéficiaires impliqués dans les efforts d'évaluation.</b>	Nous avons donné des récompenses et du soutien (bon de crédit téléphonique, frais de transport) pour motiver les bénéficiaires à participer à des enquêtes et à des séances de commentaires sur la théorie du changement, etc.	Les récompenses et le soutien de certains frais de la logistique alloués aux participants peuvent être vraiment utiles si le commissaire de l'évaluation souhaite que le processus soit aussi participatif et inclusif que possible. Cela aide à obtenir l'adhésion des différentes parties prenantes à participer, en particulier des bénéficiaires. Il faut s'assurer que ces récompenses sont distribuées par les évaluateurs ou évaluatrices et non pas par l'organisation de mise en œuvre ou le bailleur des fonds, pour éviter que ça serve comme un achat des consciences qui empêcherait une participation libre des bénéficiaires.
<b>Les informations et les coordonnées des bénéficiaires sont souvent incorrectes ou changent.</b>	Nous avons demandé au partenaire d'exécution d'organiser une réunion avec les bénéficiaires au cours de laquelle nous avons mis à jour notre liste de contacts. Nous avons eu un nombre très limité de participants et il s'agissait probablement de personnes sélectionnées par le partenaire d'exécution.	Il est important que l'organisation chargée de la mise en œuvre conserve un registre correct et à jour des bénéficiaires qu'elle sert. Un manque à ce niveau ou un enregistrement incorrect des informations sur les bénéficiaires peut entraver la participation des bénéficiaires au processus d'évaluation.

Quels défis à relever pour impliquer les parties prenantes dans la démarche évaluative	Comment avez-vous surmonté ces défis ?	Quelles leçons avez-vous tirées concernant l'inclusion des parties prenantes, en particulier des bénéficiaires, dans la démarche évaluative ?
<p><b>L'organisation de financement (bailleur de fonds) voulait que nous soumettions les rapports uniquement à eux. Nous n'avons pas été autorisés à partager les rapports avec le gouvernement, et parfois même pas avec les partenaires d'exécution.</b></p>	<p>La seule façon de partager les résultats était lors des réunions que nous avons avec les partenaires d'exécution (pendant que nous étions dans le pays par exemple) et lorsque nous étions appelés à faire des présentations lors d'ateliers ou d'événements où se trouvaient les partenaires. Nous nous sommes alors assurés d'inclure autant d'informations exploitables dans les présentations.</p>	<p>Expliquez et faites valoir aux commissaires d'évaluation que, bien que les évaluateurs ou évaluatrices soient externes, leur rôle est d'aider à l'amélioration de la mise en œuvre de l'intervention par le biais de recommandations. Une fois que les commissaires d'évaluation ont compris cela, les évaluateurs et évaluatrices peuvent s'engager plus librement avec les parties prenantes.</p>
<p><b>Parties prenantes non identifiées ou intégrées dans la conception de l'évaluation.</b></p>	<p>L'association de bénéficiaires n'ayant pas été intégrée à la phase de conception, cela ne nous a pas permis de savoir que nous aurions pu inclure les principaux bénéficiaires parmi les répondants. Au stade de la collecte des données, nous avons essayé d'improviser des entretiens avec les bénéficiaires, mais nous étions limités par le budget. Nous n'avons donc pas pu atteindre tous les bénéficiaires nécessaires. Au stade du rapport, nous avons essayé de refléter certaines contributions des principaux bénéficiaires, mais comme elles ont été exclues de la phase de conception, elles n'ont pas reçu toute l'attention qu'elles méritaient.</p>	<ul style="list-style-type: none"> <li>▪ Dès la conception, les évaluateurs et évaluatrices doivent plaider en faveur de l'identification de toutes les parties prenantes clés et de leur inclusion dans les étapes clés de l'évaluation.</li> <li>▪ Si les évaluateurs et évaluatrices remarquent qu'une partie prenante clé a été laissée de côté plus tard, mettez rapidement à jour la conception et essayez de l'intégrer de la meilleure façon possible, mais le rattrapage est souvent un défi. Il est toujours préférable de s'assurer que toutes les parties prenantes sont impliquées dès le début.</li> </ul>
<p><b>La sensibilité politique et organisationnelle empêche souvent les répondants du personnel du gouvernement/ des bénéficiaires/ des partenaires d'exécution de fournir les informations nécessaires.</b></p>	<p>La réticence des répondants à fournir des informations a rendu difficile l'obtention des informations nécessaires.</p>	<p>Un engagement précoce et spécifique à l'intervention avec les parties prenantes est critique pour identifier et définir si l'évaluation est faisable, ainsi que pour obtenir l'adhésion totale des parties prenantes clés dans le pays. Une intervention élaborée avec les parties prenantes concernées est beaucoup plus facile à évaluer, et l'engagement avec ses parties prenantes est beaucoup plus fluide, car un certain niveau d'adhésion a déjà été établi.</p>

Quels défis à relever pour impliquer les parties prenantes dans la démarche évaluative	Comment avez-vous surmonté ces défis ?	Quelles leçons avez-vous tirées concernant l'inclusion des parties prenantes, en particulier des bénéficiaires, dans la démarche évaluative ?
<b>Lorsque le personnel d'intervention a un taux de rotation élevé, cela rend difficile l'identification des parties prenantes, surtout s'il n'y a pas de mémoire institutionnelle documentée.</b>	Dans les cas où le manque de mémoire institutionnelle limitait l'accès à l'information, l'équipe a recherché les informations auprès d'autres personnes au sein de l'organisation pour aider à combler les lacunes. Lorsque cela n'était pas possible, l'absence de données a été notée.	Face au défi de la mémoire organisationnelle de l'institution, il est possible de reconstituer l'information, mais c'est un processus long et nécessitant un budget beaucoup plus élevé.
<b>Sans l'apport des bénéficiaires, il est difficile de valider les données et les rapports des exécutants et des bailleurs de fonds.</b>	Cela n'a pas été résolu et nous avons constaté l'absence de bénéficiaires et l'impossibilité de valider les données déclarées.	Face au défi de la mémoire institutionnelle/ organisationnelle, il est possible de reconstituer l'information, mais c'est un processus long et qui nécessite un budget beaucoup plus élevé.
<b>Difficulté à obtenir un accord entre le bailleur de fonds et les gouvernements.</b>	Le partenaire d'exécution n'a pas participé à la présentation des conclusions et à l'élaboration de recommandations. Nous n'avons pas surmonté l'exclusion des partenaires d'exécution.	L'exclusion d'une partie prenante importante à une étape du processus d'évaluation peut entraver et compromettre l'ensemble du processus d'évaluation. Cela peut entraîner des lacunes dans les informations et peut créer des conflits entre les autres parties prenantes.
<b>Défi d'impliquer les enfants dans la démarche évaluative.</b>	Il était important d'obtenir le consentement éclairé (y compris le consentement parental) pour tous les participants. Cela peut être un processus long et difficile. Il était également important que nous expliquions le but de l'évaluation d'une manière que les enfants puissent facilement comprendre (pendant le processus de collecte de données). L'un des moyens était de simplifier la langue, lorsqu'il s'agissait de mineurs.	Il est très important d'entreprendre une évaluation de manière collaborative. Vous devez établir des stratégies de communication claires dès le début. Cela permettra alors au processus d'évaluation de se dérouler correctement. L'institution commanditaire de l'évaluation doit être en mesure de donner aux évaluateurs et évaluatrices l'accès aux informations concernant toutes les parties prenantes concernées pour permettre de planifier et de se préparer à répondre aux exigences techniques et méthodologiques liées à l'implication de chaque type de partie prenante. Si l'évaluateur ou l'évaluatrice doit découvrir ultérieurement que les enfants font partie des répondants, cela peut affecter la démarche évaluative.

Quels défis à relever pour impliquer les parties prenantes dans la démarche évaluative	Comment avez-vous surmonté ces défis ?	Quelles leçons avez-vous tirées concernant l'inclusion des parties prenantes, en particulier des bénéficiaires, dans la démarche évaluative ?
<b>L'intérêt du gouvernement hôte et des partenaires d'exécution peut être faible, car ils perçoivent souvent l'évaluation comme une activité à cocher pour les bailleurs de fonds.</b>	Nous n'avons pas surmonté cela. La perception était si profonde que le gouvernement et les partenaires de mise en œuvre voulaient simplement se précipiter et le faire sans fournir aux évaluateurs et évaluatrices une substance solide pour soutenir la crédibilité de l'évaluation.	La participation des principales parties prenantes au processus d'intervention augmente leur volonté de participer au processus d'évaluation.
<b>Les évaluations ont souvent de petits budgets et des délais serrés, ces contraintes de coût et de temps peuvent limiter la participation des parties prenantes (et des bénéficiaires).</b>	C'est un vrai défi qui ne peut être relevé que si le bailleur de fonds comprend l'importance de l'évaluation pour l'intervention.	Si le budget et le temps sont limités, la démarche évaluative et les résultats sont très souvent un exercice de case à cocher.

## Conclusion

Bien que l'identification des parties prenantes à impliquer dans la démarche évaluative soit un processus difficile, la décision concernant l'étendue et le stade de leur implication est encore plus difficile. Cette décision est rendue plus difficile par les ressources et le temps limités disponibles pour mener les évaluations. La possibilité de marginaliser ou de favoriser certains groupes de parties prenantes, la représentation biaisée de véritables groupes de parties prenantes et l'engagement symbolique sont également d'autres facteurs auxquels les évaluateurs et évaluatrices doivent prêter attention dans leurs efforts pour impliquer les parties prenantes dans la démarche évaluative. Bien que nous convenions que l'approche d'évaluation doit être participative pour permettre une meilleure inclusion des parties prenantes dans le processus d'évaluation, dans le contexte actuel des évaluations requises dans un délai serré et un budget limité, il est important que les évaluateurs et évaluatrices ne s'enlisent pas dans des cadres théoriques, mais restent flexibles quant à la combinaison des forces des approches participatives et conventionnelles. Cela a l'avantage de permettre une participation significative des parties prenantes tout en menant l'évaluation dans un court délai et un budget limité. Le plus grand défi dans l'implication des parties prenantes dans le processus d'évaluation, cependant, est toujours au stade de développement des TdR/DP où les organisations commanditaires d'évaluation déterminent quelles parties prenantes elles souhaitent impliquer et quelle équipe d'évaluation sélectionner. Il est important pour les évaluateurs et évaluatrices ainsi que les associations d'évaluation de promouvoir et d'encourager les organisations commanditaires d'évaluation à favoriser une implication précoce de tous les représentants des parties prenantes dans le processus d'évaluation, à partir de la phase d'élaboration des TdR/DP.

## Messages clés

1. Plusieurs enjeux sous-tendent l'implication des parties prenantes dans la démarche évaluative à chacune de ses étapes de gestion. Ces enjeux tendent à compromettre l'indépendance de l'évaluateur ou de l'évaluatrice.
2. L'évaluation participative, à l'inverse de l'évaluation conventionnelle, conçoit les bénéficiaires comme des sujets plutôt que des objets d'évaluation, et permet de prendre en compte les perspectives et les aspirations de ceux-ci.
3. Il n'est pas nécessaire pour les évaluateurs et évaluatrices de se perdre dans les cadres théoriques de l'engagement des parties prenantes : il suffit de s'en inspirer et de proportionner adéquatement l'implication des parties prenantes en lien avec leur contribution.
4. Le rôle de l'évaluateur ou de l'évaluatrice dans une évaluation participative consiste non seulement à appliquer la méthodologie définie, mais aussi de veiller à la participation équitable des parties prenantes.

## Références bibliographiques

- Better Evaluation. (s. d.-a). *Étapes du processus d'évaluation*. [https://www.betterevaluation.org/fr/commissioners\\_guide\\_fr](https://www.betterevaluation.org/fr/commissioners_guide_fr)
- Better Evaluation. (s. d.-b). *Manager's Guide to Evaluation*. [https://www.betterevaluation.org/en/managers\\_guide](https://www.betterevaluation.org/en/managers_guide)
- Better Evaluation. (s. d.-c). *Understand and engage stakeholders*. [https://www.betterevaluation.org/en/rainbow\\_framework/manage/understand\\_engage\\_stakeholders](https://www.betterevaluation.org/en/rainbow_framework/manage/understand_engage_stakeholders)
- Better Evaluation. (s. d.-d). *Participatory Evaluation*. [https://www.betterevaluation.org/en/plan/approach/participatory\\_evaluation#:~:text=Participatory%20evaluation%20is%20an%20approach,the%20reporting%20of%20the%20study](https://www.betterevaluation.org/en/plan/approach/participatory_evaluation#:~:text=Participatory%20evaluation%20is%20an%20approach,the%20reporting%20of%20the%20study)
- Canadian International Development Agency. *How to Perform Evaluation : Participatory Evaluations*. <https://www.oecd.org/derec/canada/35135226.pdf>
- Centre d'excellence de la gestion axée sur les résultats. (2016). *La gestion axée sur les résultats appliquée aux programmes d'aide internationale : un guide pratique* (2<sup>e</sup> éd.). Affaires mondiales Canada. [https://www.international.gc.ca/world-monde/assets/pdfs/funding-financement/results\\_based\\_management-gestion\\_axee\\_resultats-guide-fr.pdf](https://www.international.gc.ca/world-monde/assets/pdfs/funding-financement/results_based_management-gestion_axee_resultats-guide-fr.pdf)
- Cousins, J. B. (2003). Utilization efforts of participatory evaluation. Dans T. Kellaghan, D. L. Stufflebeam (dir.), *International Handbook of Educational Evaluation*. *Kluwer International Handbooks of Education* (vol. 9, p. 245-265). Springer. [https://link.springer.com/chapter/10.1007/978-94-010-0309-4\\_16](https://link.springer.com/chapter/10.1007/978-94-010-0309-4_16)
- Cullen, A. (2009). *The Politics and Consequences of Stakeholder Participation in International Development Evaluation*. Dissertations, 656, 1-166. <https://scholarworks.wmich.edu/dissertations/656>
- Dillon, L. B. (2020). *Participatory Monitoring and Evaluation*. Sustainable Sanitation and Water Management Toolbox. <https://sswm.info/arctic-wash/module-3-health-risk-assessment/further-resources-participatory>



- Dutch Ministry of Foreign Affairs. (2009). *Evaluation policy and guidelines for evaluations*. <https://www.oecd.org/dac/evaluation/iob-evaluation-policy-and-guidelines-for-evaluations.pdf>
- ERM. (2018). *Stakeholder Engagement*. <https://www.erm.com/sustainability-report-2018/reporting-approach/stakeholder-engagement/>
- Eval, centre de ressource en évaluation. (s. d.). *Analyse des parties prenantes*. <http://www.eval.fr/methodes-et-outils/cadrelogique/analyse-des-parties-prenantes/>
- Fischler, M. (2013). *Beneficiary Assessment –a participative approach for impact assessment*. Helvetas Swiss Intercooperation. <https://www.shareweb.ch/site/Poverty-Wellbeing/Documents/Case%20M%20Fischler%20BA%20Central%20America,%20Bolivia,%20Kenya.pdf>
- Grosvenor Public Sector Advisory. (s. d.). *The key to effective stakeholder engagement*. <https://www.grosvenor.com.au/public-sector/service-delivery-improvement/key-to-effective-stakeholder-engagement/>
- Guijt, I. (2014). *Participatory Approaches*. Unicef Office of Research. [https://www.betterevaluation.org/sites/default/files/Participatory\\_Approaches\\_ENG.pdf](https://www.betterevaluation.org/sites/default/files/Participatory_Approaches_ENG.pdf)
- Guijt, I. et Gaventa, J. (1998). Participatory monitoring & evaluation : learning from change. *IDS Policy Briefing*, 12, 1-6.
- Haddaway, N. R., Kohl, C., Rebelo da Silva, N., Schiemann, J., Spök, A., Stewart, R., Sweet, J. B. et Wilhelm, R. (2017). A framework for stakeholder engagement during systematic reviews and maps in environmental management. *Environmental Evidence*, 6(11). <https://doi.org/10.1186/s13750-017-0089-8>
- Landau, P. (2019, 31 janvier). *Stakeholder vs. Shareholder : How They're Different & Why It Matters*. Project Manager. <https://www.projectmanager.com/blog/stakeholder-vs-shareholder>
- Marchand, M.-P. (2018). Où en sommes-nous avec l'implication des parties prenantes? *The Canadian Evaluation Society*, 33(2), 189-201. <https://doi.org/10.3138/cjpe.42202>
- Mcloughlin, C. et Walton, O. (2012). *Topic Guide on Measuring Results*. Governance and Social Development Resource Centre (GSDRC). <https://gsdrc.org/wp-content/uploads/2015/07/ME5.pdf>
- OCDE. (2010). *Quality Standards for Development Evaluation*. <http://dx.doi.org/10.1787/19900988>
- Patton, M. Q. et Campbell-Patton, C. E. (2021). *Utilization-Focused Evaluation*. SAGE Publications.
- Program Performance and Evaluation Office (PPEO). (2012). *Program Evaluation : Step 1 : Engage Stakeholders*. Center for Disease Control and Prevention. <https://www.cdc.gov/eval/guide/step1/index.htm>
- Salabarría-Peña, Y., Apt, B. S. et Walsh, C. M. (2017). *A Guide to the Project Management Body of Knowledge*. Project Management Institute Inc.
- Ramírez, R. et Brodhead, D. (2013). *Utilization Focused Evaluation : a primer for evaluators*. Southbound. [https://www.researchgate.net/publication/263580952\\_Utilization\\_focused\\_evaluation\\_A\\_primer\\_for\\_evaluators](https://www.researchgate.net/publication/263580952_Utilization_focused_evaluation_A_primer_for_evaluators)



- RuralInvest. (2005). *A participatory approach to identifying and preparing small scale rural investments : preparing and using project profiles*. United Nations Food and Agriculture Organization. <https://www.fao.org/3/a1420e/a1420e00.pdf>
- Shapiro, J. (2013). *Writing a Funding Proposal*. CIVICUS. <https://www.civicus.org/documents/toolkits/Writing%20a%20funding%20proposal.pdf>
- Smith, L. W. (2000, 7 septembre). *Stakeholder analysis : a pivotal practice of successful projects*. [symposium]. Project Management Institute Annual Seminars & Symposium, Houston, Texas. <https://www.pmi.org/learning/library/stakeholder-analysis-pivotal-practice-projects-8905>
- United Nations Development Program. (s. d). *Participatory Evaluation*. <http://web.undp.org/evaluation/documents/whop2.htm>
- Vital Voices Global Partnership. (2017). *Tracking Impact : A Handbook for High Quality Data Collection*. <https://www.awid.org/sites/default/files/atoms/files/handbook-final-single20page-2-11.pdf>
- Weiss, C. H. (1993). Where politics and evaluation research meet. *American Journal of Evaluation*, 14(1), 93-106. <https://doi.org/10.1177%2F109821409301400119>
- World Bank. (1996). *The World Bank participation sourcebook*. <https://doi.org/10.1596/0-8213-3558-8>
- World Vision : Strategy and Evidence Unit. (2015). *Typology of beneficiaries*. [https://www.wvevidence4change.org/wp-content/uploads/Typology-of-beneficiaries\\_full\\_v1.pdf](https://www.wvevidence4change.org/wp-content/uploads/Typology-of-beneficiaries_full_v1.pdf)
- Zukoski, A. et Luluquisen, M. (2002). Participatory Evaluation What is it? Why do it? What are the challenges? *Community Based Public Health Policy Pract.*, 5, 1-6. <https://pubmed.ncbi.nlm.nih.gov/11982021/>

# Participation des communautés locales, facteur de succès dans les évaluations adaptées à l'Afrique

*Anna Gueye*

*N'Diamé Gueye*

## Introduction

Dans le contexte de l'aide au développement, la Déclaration de Paris sur l'efficacité de l'aide au développement en 2005 met en lumière l'importance du rôle de l'évaluation dans les décisions politiques et la responsabilisation partagée des partenaires du développement. Pourtant, il est encore souvent déploré que l'exercice de l'évaluation soit réalisé selon des lignes directrices et des méthodologies prédéfinies par les bailleurs. La révision des critères d'évaluation de l'Organisation de coopération et de développement économique (OCDE) en 2019 en fut un bon exemple d'ailleurs. À cet égard, il faut se demander si la réalisation d'une évaluation doit se résumer à suivre une *liste de contrôle* de questions ou de critères afin de juger de sa crédibilité, de sa qualité ou de sa pertinence. La pluralité des contextes des pays d'Afrique et des Caraïbes ne suggère-t-elle pas de s'interroger sur la latitude des professionnels de l'évaluation à exercer un tel leadership ?

Généralement, les évaluations sont basées sur des méthodes uniformisées et conformes aux normes de l'OCDE. Or, de plus en plus, ce type d'approche est largement décrié dans le domaine comme étant inapproprié, voire non adapté au contexte des pays en voie de développement. Ces derniers militent habituellement en faveur de méthodes pertinentes, efficaces et adaptées, qui tiennent compte des valeurs, des suppositions et des pratiques locales, en un mot, une « évaluation adaptée à l'Afrique » (BAD/IDEV, 2019). Dans cette foulée, les orientations de pratiques d'évaluation adaptées à l'Afrique suggèrent, entre autres, un partenariat avec les communautés locales dans les évaluations. Or, du fait que la plupart des évaluations sont commanditées par des bailleurs de fonds, le partenariat semble illusoire (Frehiwot, 2019).

Ce chapitre s'intéresse à l'importance de l'implication des communautés locales dans une démarche d'évaluation et propose des pistes de solutions qui pourraient la favoriser.

## **Vers un partenariat des communautés locales : le renforcement des capacités des acteurs**

Au cours des dernières années, le modèle d'évaluation adaptée à l'Afrique (EAA) (mouvement « Made in Africa ») suscite un intérêt grandissant dans une perspective d'amélioration de la fonction évaluative. Le forum de Bellagio, en 2012, sur « l'évaluation proprement africaine » ou celui axé sur l'Afrique à Chilisa, en 2015, et les nombreuses initiatives (Baraza en Ouganda, CLEAR), en sont de bonnes illustrations. La notion d'EAA mise sur une approche endogène dont les bases conceptuelles tiennent compte des valeurs du continent. Dans cette foulée, le partenariat de toutes les parties prenantes, notamment les communautés locales, s'avère porteur de succès.

Une des conditions de réussite de ce partenariat est la capacité des acteurs à s'impliquer activement. La capacité se traduit par l'aptitude à concevoir et à mettre en œuvre efficacement des actions planifiées pour atteindre un objectif commun. Elle se mesure par les résultats obtenus en biens et services et en changement de comportements de la population ciblée.

Une approche basée sur le partenariat avec les communautés locales devrait se concentrer sur des stratégies d'implication des communautés et d'un référentiel de compétences de ces derniers. Des référentiels et modèles existants sont très répandus et peuvent servir de référence : l'EFQM1 (*European Framework for Quality Management*); le *World Alliance for Citizen Participation* (CIVICUS), une alliance internationale d'organisations de la société civile; le concept Baraza, mécanisme de contrôle et d'engagement communautaires; le Modèle CLEAR de participation citoyenne dans les administrations locales.

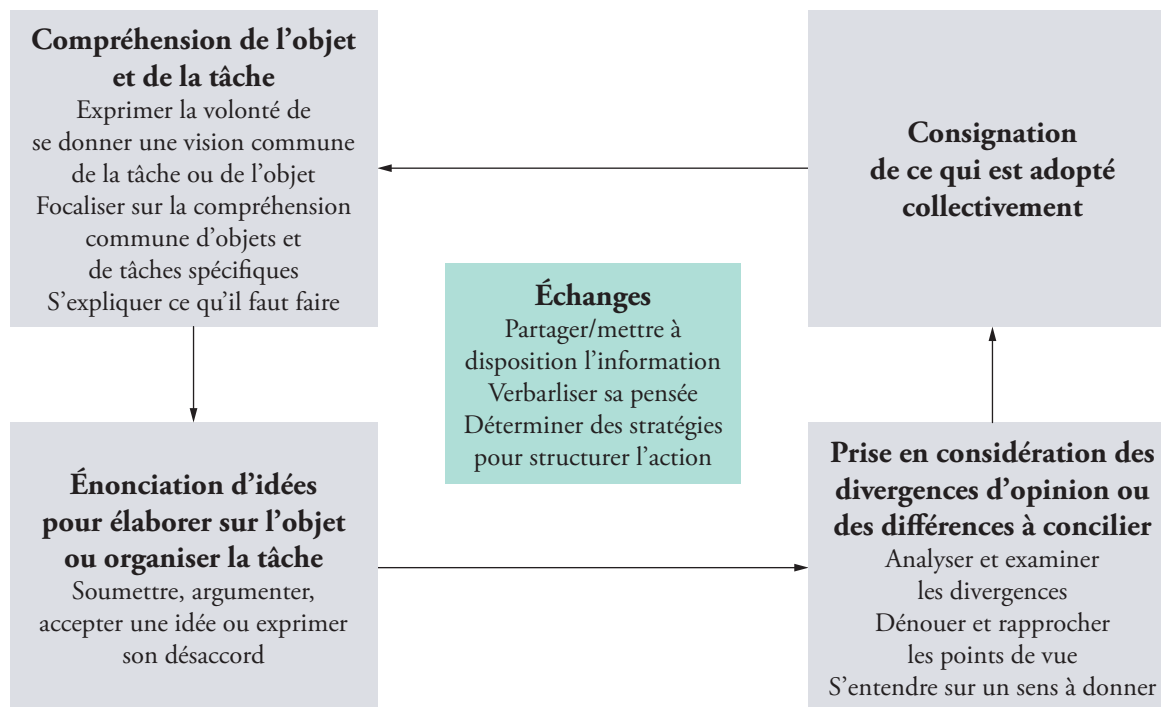
## **Pour une vision commune des pratiques d'évaluation : la négociation de sens**

Il est souvent décrié que le faible taux d'alphabétisme des communautés constitue une entrave à leur implication et à leur compréhension des grands enjeux de l'évaluation (Barry, 2004). Or, la perspective d'une vision commune qui, elle aussi, passe par une compréhension commune des objets de l'évaluation, de leur pertinence ou de leur utilité à servir l'intérêt collectif est à la fois un vecteur de mobilisation communautaire et en fin de compte une condition sine qua non pour atteindre les résultats souhaités.

Nous proposons la négociation de sens (voir schéma), comme un modèle des interactions sociales parce qu'elle comporte un versant discursif, qui met l'accent sur la communication orale, et un versant développemental, qui renvoie tant aux effets sur les individus qu'aux artefacts produits. De fait, la négociation de sens présume la conclusion d'un accord, comme le fait de négocier un prix, et une adaptation graduelle, comme le fait de négocier un virage (Wenger, 2005). Son déroulement met en jeu une intersubjectivité impliquant un lieu de rencontre entre les vues des uns et des autres, une activité située qui concerne autant la tâche à réaliser ou le problème à résoudre que la diversité et la complémentarité des compétences en présence.

Accepter le principe que l'évaluation doive s'adapter aux contextes des pays du Sud (BAD/IDEV, 2019) au regard de ses approches et de ses méthodes qui sont mises en œuvre dans une perspective de quête de données, revient à dire que la prise en compte des significations locales doit figurer parmi les critères de pertinence et d'efficacité. Dans ce contexte, la négociation de sens, du fait qu'elle repose

sur un engagement actif des parties prenantes, peu importe leur niveau de décision ou leurs compétences, se révèle une stratégie pertinente à explorer dans la construction d'une vision commune des enjeux et l'atteinte des résultats souhaités.



**Figure 1**

Zones de négociation de sens.

Source : Gueye, 2015

**Zones de négociation de sens** : Les différentes zones de la négociation de sens : la compréhension de l'objet ou de la tâche, l'énonciation d'idées pour préciser l'objet ou organiser la tâche, la prise en considération des divergences d'opinions ou des différences à concilier, la consignation de ce qui est adopté collectivement. Au centre se trouve la zone des échanges qui sont des processus transversaux. Il a été développé dans le cadre de la mise en œuvre du Plan d'action pour l'amélioration du français (PAAF) à l'enseignement primaire et secondaire du ministère de l'Éducation, du Loisir et du Sport (MELS) du gouvernement du Québec. (Source : Gueye, N. 2015)

Mais alors, qu'est-ce que négocier le sens dans la perspective d'une évaluation adaptée à l'Afrique ?

La négociation de sens est, dans ce contexte, en première instance, propulsée par l'idée de se donner une vision commune négociée des enjeux de l'évaluation en promouvant des conditions favorisant l'émergence de représentations fondées sur des paradigmes afrocentriques (BAD/IDEV, 2019). Dans cette perspective, l'énonciation des idées au regard des actions à entreprendre et de la définition des éléments d'ancrage (BAD/IDEV, 2019) de l'EAA, la prise en compte des divergences locales ou des différences à concilier, la recherche et la consignation des résolutions constituent des jalons du processus.

Certes, l'EAA promeut des évaluations faites par l'Afrique en visant à accroître l'influence de méthodes, de théories et de philosophies afrocentriques à l'échelle mondiale (BAD/IDEV, 2019). Toutefois, cette idée d'une évaluation afrocentrique ne peut se concevoir dans un vacuum. Elle devrait davantage intégrer, dans sa conception, la vision jusque-là prise en compte qui constituait la base de connaissances méthodologiques et qui servait de lentilles cognitives (BAD/IDEV, 2019) dans la définition des critères d'évaluation en terre africaine, voire caraïbéenne. C'est donc dire qu'une intégration des critères ethnoafricains devra s'opérer pour une instauration optimale de l'EAA. Dans cette mouvance, la négociation de sens se révèle, dans un second niveau d'analyse, en inscrivant dans sa démarche une dimension de courtage de connaissances qui met en jeu des processus d'interprétation/traduction, de coordination et d'alignement des perspectives (Wenger, 2005) au sein de l'écosystème évaluatif, constitué de toutes les parties prenantes, notamment, de la société civile, du secteur privé, des communautés, des professionnels de l'évaluation et du système plus global (BAD/IDEV, 2019). D'ailleurs, c'est le sens du mot *adapté*.

## Conclusion

Ce chapitre se veut un plaidoyer en faveur de la contribution des communautés locales dans une démarche d'évaluation. Il repose sur la participation des communautés locales en vue de leur implication active par la négociation de sens. Celle-ci, utilisée comme toile de fond dans la démarche de l'évaluation, est une stratégie porteuse de succès, car elle contribue à la concrétisation de l'engagement des acteurs, quels que soient leur statut socio-économique et leur pouvoir au sein de la communauté.

La contribution à ce chapitre s'inscrit aussi dans la mouvance d'une évaluation adaptée à l'Afrique et la promotion de méthodes, de théories et de philosophies afrocentriques. Le renforcement des capacités et la négociation de sens constituent des stratégies d'implication des communautés au regard de l'adaptation des critères de pertinence et d'efficacité des évaluations en contexte africain.

Concrètement, adopter la négociation de sens comme modèle des interactions sociales dans le cadre d'évaluations devrait conduire à la création d'un référentiel de compétences de l'EAA. Un cahier des charges devra en démontrer l'intégration dans son offre de services et dans son plan de déploiement. De plus, des grilles d'évaluation construites sur des critères de pertinence dérivés du référentiel de compétences devraient permettre d'apprécier l'intégration des principes de l'EAA.

Entre les mains des professionnels de l'évaluation, ce sont des outils qui fournissent des jalons et des repères dans la mise en œuvre d'une démarche fondée sur l'EAA.

Par ailleurs, la négociation de sens peut bien se tenir sous l'arbre à palabres en s'inspirant de la tradition orale, parce qu'elle donne la primauté aux discussions orales pour s'exprimer sur les affaires relatives à l'évaluation! En ce lieu, le travail de courtage du professionnel intègre une perspective de renforcement des capacités et consiste à traduire et à adapter les objectifs de l'évaluation afin de rendre ceux-ci compréhensibles aux communautés locales. Au fond, c'est en créant une certaine résonance chez elles que l'on peut susciter leur adhésion et leur doter des capacités de participation active.

## Références bibliographiques

- Barry, R. (2004) *Le renforcement des capacités des communautés rurales et leur implication dans le processus de l'évaluation environnementale en Afrique francophone : cas du Burkina Faso* [Communication]. 9<sup>e</sup> Colloque international des spécialistes francophones de l'évaluation d'impact. [https://www.sifec.org/static/uploaded/Files/ressources/actes-des-colloques/ouagadougou/session-5/2\\_Barry\\_communication.pdf](https://www.sifec.org/static/uploaded/Files/ressources/actes-des-colloques/ouagadougou/session-5/2_Barry_communication.pdf)
- Évaluation indépendante du développement de la Banque Africaine de Développement (BAD/IDEV). (2019). Des évaluations adaptées à l'Afrique. Volume 1 : Approches théoriques. *eVALUation Matters*. <http://idev.afdb.org/fr/document/des-evaluations-adaptees-lafrique-volume-1-approches-theoriques>
- Frehiwot, M. (2019). L'évaluation adaptée à l'Afrique : Décoloniser l'évaluation en Afrique. Dans Évaluation indépendante du développement de la Banque Africaine de Développement (BAD/IDEV) (dir.), Des évaluations adaptées à l'Afrique. Volume 1 : Approches théoriques. *eVALUation Matters*. <http://idev.afdb.org/fr/document/des-evaluations-adaptees-lafrique-volume-1-approches-theoriques>
- Gueye, N. (2015). *Négociation de sens et conception collective d'un dispositif de développement professionnel chez des enseignants* [thèse de doctorat, Université Laval]. <https://corpus.ulaval.ca/jspui/bitstream/20.500.11794/26423/1/31719.pdf>
- Watera, J. (2019). L'indigénisation des mécanismes de S&E axés sur les citoyens : Enseignements tirés des Barazas en Ouganda. *Evaluation Matters, Banque africaine de développement (BAD)*, 2(4). [http://www.idev.afdbnet.com/sites/default/files/Evaluations/2020-03/eVALUation%20Matters%20Q4-2019-Fr\\_1.pdf](http://www.idev.afdbnet.com/sites/default/files/Evaluations/2020-03/eVALUation%20Matters%20Q4-2019-Fr_1.pdf)
- Wenger, E. (2005). *La théorie des communautés de pratique* (traduit par F. Gervais). Presses de l'Université Laval.





# L'entrée sur le terrain en évaluation : enjeux et pratiques des collectes de données



*Vénétia Sauvain*

*Papa Souleymane Koulibaly*

*Nebnoma Kader Arnaud Sawadogo*

*Maxime Desfossés*

## Introduction

Racontons une anecdote sur l'évaluation d'une intervention dont plusieurs communautés éloignées étaient bénéficiaires. Une activité officielle (indépendante de ladite évaluation) avait lieu dans la capitale et rassemblait ainsi au même endroit plusieurs représentants de ces communautés. L'occasion était belle pour collecter des informations auprès de ces chefs et conseillers sous la forme de groupes de discussion. Or, il s'est avéré que les invités pensaient que les résultats de l'évaluation leur seraient présentés ce jour-là ! Pourtant, l'institution mandante avait bien précisé l'objectif de l'événement dans sa lettre d'invitation : une collecte de données ! Malentendu de communication qui n'eut heureusement que peu de conséquences sur ladite évaluation, si ce n'est un léger malaise momentané de la part de l'équipe... Cette anecdote illustre à quel point « l'entrée sur le terrain » peut être délicate. La mise en œuvre d'une collecte de données nécessite souvent d'entrer en contact avec différentes populations et autres sources d'information. Elle constitue une étape cruciale dans toute démarche d'évaluation, car l'équipe d'évaluation<sup>1</sup> se trouve confrontée à la réalité des données disponibles et à la réceptivité des populations. Celles-ci peuvent avoir déjà vécu de mauvaises expériences d'évaluation ou de recherche, ou encore de ne pas bien comprendre ce que signifie une évaluation, et quelles conséquences celle-ci peut avoir sur leur quotidien. Alors que la solidité de la stratégie de collecte et d'analyse de données repose en partie sur le choix des méthodes et des outils, et même si l'accessibilité des populations détentrices de l'information recherchée est l'un des principaux critères de ce choix, aucune équipe ne se targuera d'avoir parfaitement envisagé toutes les possibilités et contraintes du terrain. Une collecte de données mal exécutée ou gérée de manière inefficace générera des données imprécises et de faible qualité qui ne répondront pas aux besoins de l'évaluation (Gertler *et al.*, 2011; Patton, 2014). Elle pourra donner lieu à des malentendus entre l'équipe d'évaluation et les acteurs

---

1. Nous préférons utiliser le terme « équipe d'évaluation », bien qu'il puisse s'agir d'une seule personne, un évaluateur ou une évaluatrice.

et actrices qui entourent l'intervention<sup>2</sup>. Pire encore, elle peut causer du tort aux populations. Ultimement, le niveau de préparation et la flexibilité avec laquelle l'équipe d'évaluation pourra se confronter à cette réalité du terrain seront déterminants pour juger de la réussite de la collecte. Il faut également rappeler, qu'à la différence d'un chercheur, qui a une plus grande liberté et une marge de manœuvre suffisante sur la manière de conduire sa collecte, le travail de terrain de l'équipe d'évaluation est délimité par des questions d'évaluation plus spécifiques et par la volonté des principaux utilisateurs de l'évaluation (ou d'autres parties prenantes stratégiques) de donner accès à une population plutôt qu'à une autre dans des délais courts et en fonction de leur degré de résistance à l'étude évaluative, en particulier sur les sujets difficiles.

Ce chapitre aborde les aspects pratiques à considérer pour la mise en œuvre d'une collecte de données, en particulier sur le « terrain », afin de réduire les risques de biais inhérents à celle-ci, ainsi que de faciliter ce travail de collecte. Nous avons rassemblé, à partir de nos propres expériences de terrain en Afrique et au Québec, les éléments que nous jugeons cruciaux pour l'entrée sur le terrain de la collecte. Ils peuvent être considérés comme autant de facteurs de succès ou d'échec de la collecte de données d'un point de vue opérationnel.

## Délimiter le terrain de la collecte

Il est important ici de tenter de délimiter ce que l'on entend par « terrain ». Le terme a été consacré par les champs de l'anthropologie et de l'ethnographie sous le terme « enquête de terrain<sup>3</sup> » ou « enquête ethnographique », et signifiait une visite physique sur un lieu pendant une durée assez longue avec l'utilisation de l'outil « observation », tel que vu au chapitre sur les outils de collecte (cf. encadré 1). Cependant, bien que nous mettions l'accent sur ce type de « terrain » physique dans ce chapitre, nous utiliserons le terme pour renvoyer à la collecte des données auprès de l'ensemble des sources qui constituent en elles-mêmes une unité d'analyse (c'est-à-dire des individus, des lieux ou des objets physiques, porteurs de sens). Lors de la collecte de données, les personnes qui seront sollicitées pour fournir des informations peuvent provenir de toutes les catégories d'acteurs et d'actrices de l'intervention, et leurs intérêts pour l'évaluation et leur disponibilité seront très variables. L'intérêt de ces individus pour l'équipe d'évaluation dépendra de l'implication des personnes qui les représentent dans la démarche évaluative, des relations de pouvoir ou d'autorité de ces représentants, de l'importance que ces derniers accordent à l'évaluation et de l'utilité perçue ou de leur intérêt propre vis-à-vis de ladite évaluation (Patton et Campbell-Patton, 2022). Dans ce sens, l'équipe d'évaluation peut délimiter un terrain restreint, dont les sources sont plutôt au niveau de l'autorité ou du maître d'œuvre,

2. Ce chapitre reprend les termes qui identifient les acteurs de l'intervention présentés dans le chapitre « La terminologie de l'évaluation » de Marceau et Sylvain (2022) tout en donnant des précisions lorsque nécessaire. Parfois, le terme « équipe de mise en œuvre de l'intervention » sera privilégié pour faire référence aux personnes concrètement responsables de la mise en œuvre de l'intervention sur le terrain. La notion d'autorité peut renvoyer à la fois aux bailleurs de fonds et aux ministères de tutelle des interventions de développement. Dans ce contexte, nous tenterons d'être assez précis dans nos exemples, puisque ceux-ci sont à la fois issus de nos expériences d'ONG internationales et locales au Burkina Faso et au Sénégal que d'une équipe d'évaluation externe à l'administration publique québécoise. De manière générale, nous parlerons du client principal de l'évaluation au sens de Patton (Patton et Campbell-Patton, 2022). Pour plus de références, se référer aux chapitres sur les parties prenantes, leurs implications et sur la gestion d'une évaluation.

3. Le Conseil International de la Langue Française (1976) définissait le terrain comme « l'organisation même du travail des enquêteurs » et à « la phase de recueil des informations auprès des interviewés par les enquêteurs », mais le terme « enquêteurs » fait surtout référence à des collectes de données quantitatives par questionnaire (donc des questions fermées), ce qui est limitatif dans ce contexte.

ou encore un terrain élargi qui tente d'accéder aux populations cibles de l'intervention. Et elle devra d'ailleurs être assez précise sur cette population cible : s'arrêtera-t-elle seulement aux intermédiaires (ou bénéficiaires indirects) ou ira-t-elle jusqu'aux bénéficiaires ultimes (ou bénéficiaires directs, cf. le chapitre « Les parties prenantes en évaluation : généralités ») au sens d'Affaires mondiales Canada (2016) ?

Donnons un exemple concret de « terrain ». Un ministère de l'éducation commande l'évaluation d'une nouvelle mesure contre le décrochage scolaire et informe à cet effet différentes parties prenantes en leur indiquant qu'elles seront potentiellement sollicitées dans la collecte de données : des représentants et représentantes de deux syndicats fédérateurs du personnel enseignant, des directions d'écoles et des associations fédératrices de parents d'élèves ainsi que les services déconcentrés de gestion des écoles au niveau local. L'équipe d'évaluation peut envisager un terrain assez restreint : le ministère et ses parties prenantes (au niveau de leur représentation). Cependant, elle peut aussi l'élargir au personnel de direction des écoles, au personnel enseignant, aux parents d'élèves, au personnel de soutien administratif et pédagogique, aux élèves et même à la société en général. Le terrain peut aussi être constitué d'objets physiques qui entourent l'intervention : les locaux de l'école, le matériel pédagogique, les biens immobiliers, ainsi que des objets immatériels, comme les bases de données, les documents de procédure, etc.

Étant donné l'avancée fulgurante des moyens de communication à distance, dans les sociétés modernes, nous considérons que l'expression « entrée sur le terrain » ne veut pas nécessairement dire être présent physiquement sur le terrain. S'il s'agit d'un sondage en ligne auprès du personnel enseignant, cela reste que l'équipe d'évaluation devra réfléchir à la manière dont les enseignantes et enseignants seront sollicités pour répondre au sondage. Dans notre pratique au Québec, nous avons vu que les directeurs d'écoles, eux-mêmes informés par les intermédiaires gouvernementaux ou par une directive du ministère de l'Éducation, informeront le personnel enseignant de la tenue d'une collecte de données par sondage électronique. Même si l'entrée sur le terrain n'est pas physique, la série d'éléments pratiques présentés dans ce chapitre sont importants pour la réussite de la collecte.

### **Encadré 1 – Origine de l'expression « enquête de terrain » ou « enquête ethnographique »**

Les origines de l'enquête de terrain (ou enquête ethnographique) peuvent être attribuées à la fois aux travaux anthropologiques de Malinowski et à l'École de Chicago. Dans une perspective traditionnelle, l'étude des populations indigènes des îles Trobriand par Malinowski, en 1914, marque le début des études des autres cultures, qui, selon une perspective occidentale, paraissent radicalement différentes, par la mise en œuvre de la première observation participante (Marchive, 2012). Ces études s'inscrivaient dans une perspective coloniale (Blanc, 2012; Given, 2008). Puis, l'enquête de terrain est marquée par la « fin de l'exotisme » et évolue vers l'anthropologisation de tout ce qui entoure le quotidien (Bensa, 2006), c'est-à-dire par « l'ampleur des réflexions sur les façons de l'effectuer et de le penser » (Blanc, 2012). En parallèle, il s'entend une décolonisation de la recherche; les rapports de pouvoir entre le chercheur et les personnes, objets d'étude, sont déconstruits. Une certaine réflexivité du chercheur sur lui-même (quels sont les biais qui lui sont propres, issus notamment de ses référents culturels, et qui structurent sa perception de ce qui est « autre » que soi-même) est préconisée, ainsi que la participation des individus, qui sont l'objet de la recherche, à la recherche elle-même, comme acteurs et actrices (Laurent, 2009; Salün, 2012). L'une des critiques faites par Edward Saïd, à l'anthropologie en général, renvoie à la façon dont les savoirs occidentaux créent des stéréotypes sur l'Orient.

Ceux-ci tendaient à créer un « Autre », « reposant sur l'incapacité supposée [des groupes étudiés] [...] à se représenter eux-mêmes » (Weber, 2015). L'autre origine de l'expression « enquête de terrain » peut être attribuée au courant de pensée sociologique américain, l'École de Chicago. En souhaitant étudier directement l'occupation des espaces urbains par les populations immigrantes selon une approche compréhensive en évitant les statistiques et autres instruments de mesure quantitatifs (Dufour *et al.*, 1991), elle s'inscrit dans une perspective empirique de la pensée sociologique du début du 20<sup>e</sup> siècle (Blanc, 2012; Marchive, 2012). L'observation est alors particulièrement utilisée. Elle l'est encore aujourd'hui, par l'enquête ethnographique; celle-ci demeurant un « moyen incontournable d'accès à la connaissance dans le champ des sciences sociales » (Marchive, 2012). Ainsi, il est intéressant pour les évaluateurs et évaluatrices de bien comprendre d'où viennent les courants de pensée dans lesquels leurs outils de collecte de données s'inscrivent. Il est utile de rappeler que bien des discussions ont eu et ont encore cours dans le domaine des sciences sociales sur l'application des techniques de collecte de données dans une perspective de décolonisation des savoirs. Rien ne sert aux évaluatrices et évaluateurs de réinventer la roue...

## Préparer son entrée sur le terrain

Préparer son entrée sur le terrain renvoie entre autres à la formation et à la gestion des ressources (humaines, matérielles et financières) mobilisées pour les activités de collecte, à la stratégie de communication auprès des actrices et acteurs auprès desquels l'équipe d'évaluation souhaite recueillir de l'information, à l'identification des intermédiaires ou des personnes-ressources pour y accéder, ainsi qu'à la prise en compte des caractéristiques socioculturelles propres à chaque contexte. Les efforts techniques, matériels et financiers de cette préparation sont non négligeables. Des modifications pourront être apportées à la matrice d'évaluation face à la réalité du terrain. Comme l'occurrence et l'importance des éléments contextuels peuvent être variables d'une évaluation à une autre, il est hasardeux d'établir une liste de contrôle applicable à toutes les situations. Il n'en demeure pas moins que plusieurs aspects peuvent faire l'objet d'une réflexion collective et approfondie. Nous vous présentons les éléments les plus saillants.

## Répercussions de l'ancrage institutionnel de l'évaluation et de l'inclusion des acteurs et actrices sur la collecte

L'information collectée sur le terrain a une valeur, et l'entrée sur le terrain peut être envisagée comme un processus de négociation avec ses usages et sa diplomatie qui lui sont propres. Que souhaite obtenir l'équipe d'évaluation? Une certaine adhésion des personnes sollicitées pour que l'information collectée contribue au respect des critères de qualité présentés dans le chapitre sur les préalables méthodologiques.

Il arrive parfois que les activités de collecte de données se heurtent à de l'incompréhension ou à des réticences de la part des personnes chez qui l'information doit être récoltée. Soit que le mot « évaluation » fasse peur, soit que les communautés aient déjà eu trop souvent à faire avec des collectes de données avec toutes sortes de finalités (par des équipes d'enquête pour obtenir les données de la situation de départ, des équipes de recherche, des équipes d'évaluation, etc.) Ces situations ont été observées dans différents contextes par les auteurs. Certains sujets potentiellement délicats (travail

des enfants, violences faites aux femmes et aux filles, santé sexuelle et reproductive) peuvent également créer des réticences. D'après nos expériences, ces réactions peuvent conduire à un faible niveau de participation, à de la rétention d'informations, à la distorsion des faits rapportés, à de nombreux cas de non-réponses ainsi qu'à divers biais non liés à l'échantillonnage ou à l'instrument de mesure (erreurs de mesures, biais de désirabilité sociale, etc.). Il est indispensable que l'équipe d'évaluation identifie préalablement ces risques, en tienne compte et adopte les stratégies appropriées pour les prévenir. Plusieurs approches de mitigation peuvent être employées à cet effet. Nous retenons deux axes de préparation en amont de la collecte : le lien avec les populations sollicitées et la préparation des membres de l'équipe d'évaluation dédiés à la collecte.

### **Préparer l'entrée sur le terrain en amont : le lien avec les populations sollicitées**

Dans la stratégie de communication générale prévue pour l'évaluation, une réflexion sur la manière dont seront sollicitées les populations concernées par la collecte des données doit être prévue. Une communication proactive auprès des représentantes et représentants clés des populations concernées par les opérations de collecte de données, que Patton désigne par le terme « gardien du temple » (*gatekeeper*) permettra de démystifier les objectifs de l'évaluation et de favoriser une meilleure compréhension des besoins de la collecte d'information. Le but est de susciter la confiance envers l'équipe d'évaluation, de dissiper les craintes liées à la démarche et de favoriser la collaboration d'un maximum de personnes. Qui est responsable de prévenir les populations concernées qu'une collecte de données va avoir lieu dans le cadre d'une évaluation ? L'équipe d'évaluation ? Le client principal de l'évaluation ? L'équipe de mise en œuvre de l'intervention qui est constamment sur le terrain ? Comment cette communication aura-t-elle lieu ? Qui participera à la collecte ? Généralement, et sans même parler d'approches participatives, ces questions doivent minimalement avoir été clarifiées avant même d'informer la population concernée (voir le chapitre « La gestion d'une évaluation » concernant notamment la mise en place d'un comité d'évaluation à qui ces questions peuvent être soumises). Contrairement au contexte d'une recherche, où le chercheur décide seul de la manière dont il veut procéder, l'équipe d'évaluation devra prendre en compte la perspective du client principal de l'évaluation (Patton, 2014). De plus, l'inclusion d'autres parties prenantes stratégiques pour le succès de la collecte de données bien en amont – par exemple au niveau des questions d'évaluation, des méthodologies, des choix et de l'élaboration des outils de collecte et du calendrier – constitue également un bon moyen de favoriser une meilleure adhésion des personnes sollicitées par la collecte de données (Jacob et Ouvrard, 2009). La connaissance du milieu par les actrices et acteurs locaux permettra également d'anticiper certains irritants qui pourraient se présenter et d'éviter des plaintes en rapport avec les activités de collecte. Enfin, la participation de ces actrices et acteurs pourra en faire de précieux alliés si des difficultés se présentent ou si un plaidoyer en faveur de la poursuite des activités de collecte doit être mené, notamment auprès des autorités coutumières, dans des concessions ou dans certains groupements. De plus, il est essentiel de bien savoir quelles informations seront communiquées concernant la finalité de la collecte. Cette présentation doit être bien réfléchiée en amont entre l'équipe d'évaluation et le comité d'évaluation (en raison du mot « évaluation » qui peut susciter des méfiances comme nous l'avons déjà expliqué). Plusieurs cas de figure sont possibles. En tant qu'équipe d'évaluation d'une organisation externe, pour donner le plus possible la perception que l'évaluation est « indépendante », l'équipe préférera être seule à collecter l'information sur le terrain (sans représentant du client principal notamment). En fonction de ce qu'il aura été convenu, les populations seront informées de la venue de l'équipe d'évaluation sur le terrain soit par le client ou par l'équipe d'évaluation elle-même. Parfois, lorsque l'intervention est mise en œuvre par une organisation maître

d'œuvre ou exécutante très impliquée sur le terrain, il peut être judicieux que ce soit cette organisation qui informe les populations sur place (plutôt que le client principal de l'évaluation). Lorsqu'il s'agit de l'évaluation d'un programme ou d'un service public sous la responsabilité d'un ministère sectoriel, l'institution demandeuse de l'évaluation, souvent un organisme central (par exemple, une Cour des comptes, ou encore l'équivalent du Secrétariat du Conseil du Trésor ou d'un ministère qui octroie des crédits, etc.) informera elle-même les ministères sectoriels de la venue d'une équipe d'évaluation (interne ou externe). Parfois, l'équipe d'évaluation peut avoir organisé des ateliers préliminaires d'information visant à expliquer aux actrices et acteurs locaux et aux représentantes et représentants des populations desservies par le projet les tenants et aboutissants de la collecte de données sur divers aspects (nature des informations recherchées, calendrier de collecte, gestion des plaintes, gestion de la confidentialité et garantie de l'anonymat, utilisation des résultats). Ces moments de préparation sont également des occasions d'échanges en vue de faire état des préoccupations et des questionnements venant des participants. Ils permettent à cet effet d'apporter des clarifications et de les rassurer sur la démarche. Les participants à ces rencontres pourront à leur tour faire office de relais au sein des populations ciblées pour la collecte de données et ainsi mieux les y préparer. Cette approche a été notamment utilisée dans le cadre d'une expérience pilote de promotion des pratiques d'hygiène chez des élèves du primaire au Burkina Faso. Les rencontres préliminaires d'information regroupant du personnel de direction des écoles, du personnel enseignant, des inspecteurs et inspectrices en éducation, des parents d'élèves et du personnel infirmier ont permis de décrire les éléments caractéristiques de l'intervention, de répondre à différentes questions et de démystifier la démarche. Cette façon de faire a été largement appréciée par les participants et a grandement facilité leur collaboration pour l'ensemble de la mise en œuvre de l'intervention.

En fonction des contextes, le client principal, l'équipe de mise en œuvre de l'intervention ou l'équipe d'évaluation pourrait avoir à formuler des demandes d'autorisation adressées à des structures administratives et policières locales avant de lancer les opérations de collecte de données dans une localité. Elles ont pour but d'informer ces autorités sur les objectifs de la démarche, la durée des activités, l'identité des personnes qui seront sur le terrain, les sites visités et parfois l'utilisation des résultats. Le temps pour obtenir cette autorisation doit être prévu dans la démarche de collecte. À titre d'exemple, dans le cadre d'un projet visant à lutter contre la propagation d'une maladie infectieuse au Burkina Faso, l'équipe d'évaluation a dû rédiger une lettre informant les autorités administratives locales (directions régionale et provinciale de la santé) des opérations de collecte de données qu'elle s'apprêtait à faire.

Les autorités coutumières ne sont pas en reste. Bien qu'elles n'exercent pas de pouvoir formel au même titre que les structures administratives, notre expérience a montré qu'une relation de confiance avec les populations et les acteurs locaux avant d'entamer des travaux avec eux renforce la crédibilité, la fiabilité et la confirmabilité des données collectées. L'engagement des notables communautaires et des personnes-ressources s'avère souvent essentiel pour convaincre les populations et les mettre en confiance. Leur soutien est crucial, car ils sont souvent en mesure d'influencer les attitudes des membres des communautés par rapport à l'enquête et leur disponibilité à fournir des informations dans les délais requis. La compréhension par les acteurs locaux des objectifs de la collecte et du type d'informations que l'équipe d'évaluation souhaite recueillir permettra de mieux cibler les personnes à interviewer ou les endroits à visiter (expérience et vécu par rapport à l'intervention ou par rapport au domaine d'intervention) et de s'assurer d'une certaine représentativité et d'une diversité des populations (âge, genre, niveau de revenus, etc.).



Enfin, la nécessité d'obtenir le consentement préalable de ceux et celles qui fourniront les informations, le respect des législations des pays en matière de confidentialité des données, le respect de l'anonymat, sont autant de principes éthiques qui doivent être préparés avant l'entrée sur le terrain. Ce détail demande une attention particulière, car des enjeux importants d'un point de vue culturel ne sont pas toujours maîtrisés par les consultants éloignés de ces réalités, et peuvent entraver la participation des différents groupes sociaux aux collectes.

L'ensemble de ces démarches de communication préalables est en lien direct avec la gestion du temps de collecte de données. Le temps nécessaire dépendra de la portée de la collecte et de la disponibilité des personnes à rencontrer. S'assurer de la disponibilité de ces dernières, par exemple, prévoir un temps suffisant pour interagir avec l'équipe d'évaluation lors de visites dans leur communauté ou pour donner le temps aux individus de répondre à un sondage en ligne (généralement de l'ordre de trois semaines). Notre expérience nous a montré que le calendrier des populations cibles ne coïncide pas souvent avec l'agenda de l'évaluation. D'après nos expériences, on peut citer comme facteurs l'interférence des calendriers agricoles (transhumance, récoltes, etc.), les enjeux politiques (élections), les aspects socioculturels et religieux (jours fériés, mariage, etc.) qui seront toujours prioritaires pour les populations par rapport à une collecte de données pour une évaluation (cf. encadré 2).

### Encadré 2 – Concilier le temps de l'évaluation avec le temps des populations

Dans le cadre d'un programme de renforcement des capacités sur la résilience et l'adaptation aux changements climatiques (*Building Resilience and Adaptation to Climate Extremes and Disasters*, BRACED)<sup>4</sup>, le protocole d'enquête a été élargi pour couvrir davantage de villages que prévu pour une évaluation des effets comportant une mesure avant/après. Comme plusieurs villages ont été intégrés au cours de la mise en œuvre du programme, et qu'ils n'étaient pas inclus dans les données de la situation de départ (*baseline*), l'équipe a tenté d'utiliser des questions de rappel pour permettre de reconstruire celle-ci. Or, malgré les bonnes dispositions prises par l'équipe du programme en termes de communication avec les différentes parties prenantes, et bien que ces dernières aient validé la période d'enquête initialement envisagée, celle-ci a dû être décalée d'un mois en raison du calendrier agricole des paysans. En effet, la première moulture coïncidait avec la récolte du mil, d'autres céréales ainsi qu'avec le dessouchage des premières variétés d'arachides et d'oléagineux, ce qui rendait les paysans indisponibles pour la collecte.

### Préparer l'entrée sur le terrain en amont : les personnes responsables de collecter des données

D'après notre pratique, on retrouve deux cas de figure de collecte de données : des situations où des évaluatrices et évaluateurs « de métier », et parfois leurs équipes (que ce soient des experts sectoriels ou du personnel assistant), sont les personnes qui vont directement sur le terrain pour mener les

4. Programme financé par la coopération britannique de 2015 à 2019, dans les départements de Kaffrine, Koungeul, Malem Hodar et Birkelane au Sénégal, avec l'appui de Innovation Environnement Développement Afrique (IED Afrique), sur l'accès et la gestion de ressources en lien avec l'adaptation locale aux changements climatiques ainsi que sur l'intégration de la résilience climatique au sein de leurs systèmes de planification et de budgétisation : <http://www.braced.org>



collectes; ou des situations où sont recrutés des « enquêteurs » et « enquêtrices », parfois appelés facilitateurs ou facilitatrices, encadrés par des « superviseurs » (l'ensemble formant ce qu'on appelle dans ce chapitre l'équipe d'enquête). Ce dernier cas de figure a plutôt lieu lorsque l'équipe d'évaluation souhaite envoyer sur le terrain une équipe d'enquête pour la réalisation de sondages ou de questionnaires (donc des outils quantitatifs où la majorité des questions sont fermées, ou alors plus rarement des entrevues structurées). La qualité des informations obtenues dans ce cas dépend en grande partie du travail effectué par ces équipes déléguées sur le terrain. L'expérience des auteurs du Burkina Faso et du Sénégal leur a montré qu'il est important de sélectionner l'équipe d'enquête en fonction de sa capacité et de sa motivation à réaliser les tâches de manière rigoureuse et parfois protocolaire (cf. les encadrés 3 et 4). La capacité à suivre les instructions de l'enquête et à les appliquer de manière précise et exacte, à savoir se présenter et présenter la raison de la collecte, la politesse et l'entregent avec les populations, la maîtrise de la langue de la communauté, sont autant de compétences et de savoir-être recherchés auprès des personnes envoyées sur le terrain. La préparation de l'équipe d'enquête inclut une bonne compréhension des informations nécessaires pour l'évaluation, du contexte de l'évaluation et des populations avec qui elle aura à interagir. Le cas échéant, l'équipe d'évaluation peut former l'équipe et mettre à sa disposition un manuel de procédures ou de référence à des fins d'encadrement, de renforcement des capacités et d'uniformisation des techniques de collecte de données quantitatives (cf. encadré 5). Cette période de formation ne doit pas être sous-estimée et elle prend du temps.

### **Encadré 3 – Quelles tâches sont le plus couramment réparties entre les superviseurs et les enquêteurs sur le terrain ?**

Pour les superviseurs :

- Suivre la manière dont les enquêteurs effectuent leur travail;
- S'assurer que l'ensemble des instructions sont prises en compte par les enquêteurs;
- Répondre aux questions que se posent les enquêteurs;
- Contrôler la qualité des données collectées;
- Servir d'intermédiaires entre l'équipe de projet et les enquêteurs.

Pour les enquêteurs :

- Identifier les ménages ou les personnes spécifiques qui font l'objet de la collecte selon le protocole prévu;
- Obtenir le consentement des personnes auprès desquelles sera soumis le questionnaire;
- Faire remplir le questionnaire selon le protocole prévu;
- Enregistrer les réponses selon le protocole prévu.

Idéalement, les superviseurs auront une expérience antérieure de terrain en tant qu'enquêteurs dans des enquêtes ayant été bien conduites. La pratique a montré qu'il faut sélectionner toujours plus d'enquêteurs que nécessaire, les former et choisir les meilleurs après le *coaching*. Une liste d'attente peut être prévue pour pallier les désistements. Lorsqu'une équipe d'enquêtrices et d'enquêteurs est mise en place, l'inexpérience n'est pas nécessairement un désavantage. En effet, l'exemple présenté dans l'encadré 4 illustre à quel point les mauvaises habitudes de travail d'un enquêteur, malgré son expérience, peuvent porter préjudice à l'enquête. Le respect des instructions édictées étant un élément essentiel de contrôle de qualité pour l'équipe d'évaluation. La volonté de suivre sa propre logique est un trait de caractère à éviter, quel que soit le niveau d'expérience. L'implication d'enquêtrices et d'enquêteurs expérimentés pour construire le protocole et le questionnaire peut en revanche constituer un atout.

#### **Encadré 4 – Choisir un enquêteur expérimenté représente-t-il nécessairement un avantage ?**

En 2010, un vaste programme de revalorisation des chaînes de valeurs Mangue et Anacarde a été financé sur cinq ans au Sénégal, en Gambie et en Guinée-Bissau dans le cadre de la lutte contre la pauvreté. Au Sénégal, le programme intervenait dans toutes les régions de la Casamance et dans le département de Foundiougne situé au centre, dans la région de Fatick. En 2013, lors de l'élaboration de la situation de référence, dix équipes d'enquêteurs ont été formées pour couvrir les dix départements d'interventions du programme. Seule l'équipe de Foundiougne était constituée de jeunes enquêteurs; les neuf autres se distinguaient par le caractère expérimenté de ses agents, composées du personnel professionnel du développement émanant des services de l'État (agriculture, élevage, pêche, etc.) et du secteur privé. À l'issue des enquêtes, les données d'un des départements ont été rejetées, car il s'est avéré que les enquêteurs partageaient des données fictives du fait qu'ils avaient une relative maîtrise de la zone et que la collecte de ce type de données faisait partie de leur travail quotidien (mission régaliennne de leurs services d'appartenance). Pour ceux d'entre eux qui faisaient l'effort du déplacement sur le terrain, ils remplissaient les bases selon leur propre compréhension de la situation, en faisant fi des orientations du projet. Près de trois mois après la fin des études de terrain, les populations de certains villages sont entrées en contact avec l'équipe du projet pour s'informer de la situation des enquêtes : ils s'indignaient de ne pas participer aux enquêtes alors qu'ils faisaient partie de l'échantillon. L'équipe de projet a ainsi pris conscience du caractère fictif des données partagées par certains enquêteurs, et, par la même occasion, du manque d'implication des superviseurs sur le suivi des informations collectées. Une reprise des travaux de collecte dans ce département a dû être prévue. L'équipe du projet a alors placé sa confiance en ses jeunes enquêteurs inexpérimentés, initialement déployés dans le département de Foundiougne, en dégageant des ressources supplémentaires pour la logistique et les honoraires. L'exercice a été un succès qui a abouti au recrutement de deux agents parmi les jeunes enquêteurs inexpérimentés.

### **Encadré 5 – Que peut comprendre un manuel de procédures pour l'équipe d'enquête ? (Source : Statistique Canada, 2010)**

- Une section, assez générale, qui rappelle à l'enquêteur ou au facilitateur les objectifs de l'évaluation, de la collecte de données et éventuellement du type d'analyse qui sera réalisée à partir des données colligées;
- Une section précisant la façon d'aller à la rencontre des répondants. Cette partie contient également un texte de présentation que l'enquêteur devra lire aux personnes rencontrées. Le texte offre notamment une description du projet, de l'organisme chargé de sa mise en œuvre, de l'organisation chargée de l'évaluation si celle-ci diffère de l'organisation chargée de la mise en œuvre et de l'objectif de l'enquête ou de l'entrevue. Il y est également fait mention de la garantie de la préservation de la confidentialité et du respect de l'anonymat des personnes rencontrées. Il se termine par une question visant à obtenir l'accord de participer à l'enquête ou à l'entrevue;
- Le questionnaire d'enquête (support d'un sondage) ou le guide d'entretien (support de l'entrevue structurée);
- Pour les questionnaires/sondages : une section décrivant la marche à suivre en cas de non-réponse ou d'attrition (nombre de tentatives pour obtenir la participation des personnes interpellées, procédure de remplacement, etc.);
- Une section permettant à l'enquêteur de rendre compte du déroulement de ses activités et de ses observations;
- Dans les techniques qualitatives d'observation, il est habituel de tenir un « journal de bord » où seront consignées les observations de la journée ainsi que les réflexions de l'enquêteur ou enquêtrice.

Que ce soit en zones rurales, périurbaines ou urbaines, il y a de fortes chances que la collecte de données se fasse auprès de populations dont la langue d'usage diffère de celle avec laquelle les instruments de collecte de données ont été élaborés (généralement en français ou en anglais). De plus, de nombreux interlocuteurs peuvent démontrer un niveau variable d'aisance dans une langue ou dans une autre. Ce décalage peut avoir des conséquences directes sur la validité des données colligées entraînant de nombreux biais (erreurs de mesure, non-réponses, distorsion) du fait des risques d'incompréhension entre l'équipe d'enquête et les répondants. L'équipe d'évaluation devra s'assurer, au moment d'engager les personnes qui réaliseront la collecte de données sur le terrain, que ces dernières ont une connaissance suffisante des principales langues utilisées dans les régions ciblées. Ces personnes devront également avoir une bonne maîtrise de la langue employée dans les outils de collecte de données afin de traduire de façon précise les questions posées et les réponses obtenues. Dans la mise en œuvre des outils qualitatifs, il peut être également intéressant de former les personnes-ressources/intermédiaires sur le terrain qui maîtrisent au mieux la langue de la capitale et la langue locale dans le but qu'elles collectent elles-mêmes les données.

Après le recrutement, une période de rodage est à prévoir quelques jours avant le lancement des activités de collecte pour permettre l'appropriation des outils par l'équipe d'enquête. Cette période offre également l'occasion de vérifier la correspondance de certains termes en français ou en anglais

avec la langue locale. Elle permet ainsi d'apporter des correctifs aux instruments avant leur utilisation. Le rodage se fait généralement lors d'une ou de plusieurs rencontres lors desquelles l'administration des outils est simulée à travers des mises en situation et des jeux de rôle. Un prétest peut également être réalisé auprès de quelques personnes représentatives des populations ciblées afin d'avoir une bonne idée de l'adéquation des termes traduits.

Enfin, l'encadré 6 donne un aperçu d'autres éléments logistiques à prendre en compte dans la préparation de la collecte de données.

### **Encadré 6 – Autres aspects logistiques à considérer**

La préparation de la collecte de données requiert de prévoir la logistique des éléments suivants :

- Le matériel bureautique et documentaire en nombre suffisant pour l'ensemble de l'équipe d'enquête (stylos, crayons, questionnaires papier, guides d'entretien);
- Les tablettes en nombre suffisant si la collecte de données se fait électroniquement (avec la possibilité de synchroniser les données au besoin);
- Le matériel d'enregistrement pour la collecte de données qualitatives, le cas échéant;
- Des téléphones pour les enquêteurs et les superviseurs, ou des crédits de communications (prépayés) pour favoriser un contact constant avec le terrain;
- Des bons d'essence ou des montants d'argent destinés aux déplacements sur le terrain;
- Un lieu sécuritaire d'entreposage pour les données collectées sous format papier;
- La logistique en lien avec la sécurité des personnes qui se déplacent sur le terrain, le cas échéant.

## **Être sur le terrain : une articulation entre les procédés et techniques quantitatives par rapport aux techniques qualitatives.**

Le respect des critères de qualité des méthodes quantitatives et qualitatives implique parfois des procédés et des techniques de mise en œuvre à l'opposé les uns des autres. Par exemple, un sondage en présentiel ou au téléphone nécessitera une constance dans l'administration des questions fermées, sans que les enquêteurs et enquêtrices procèdent à des reformulations ou à des explications de ces questions; ou encore sans qu'ils puissent répondre à la place des répondants en interprétant un silence pour un « je ne sais pas » ou un simple oui pour un « tout à fait satisfait ». Plusieurs techniques de contrôle de qualité pour l'administration de ces questionnaires existent et sont généralement explicitées dans le manuel de procédures (cf. encadré 5). À l'inverse, les techniques d'entrevues, particulièrement celles dites libres, et l'observation participante, favoriseront les relations et les conversations informelles avec les participants dans le but de gagner la confiance des individus et de générer un maximum d'informations. La notion de neutralité empathique est alors privilégiée. Celle-ci fait référence à l'attitude de la personne qui réalise l'entrevue (ou qui met en œuvre un outil de collecte

participatif) qui doit conserver à la fois une attitude neutre (ne pas émettre son opinion sur le sujet abordé) et empathique (montrer que l'on écoute attentivement son interlocuteur, qu'on le comprend...). En ayant une connaissance fine des procédés et des techniques préconisées pour chaque outil, couplé à une réflexion approfondie de la réalité du terrain, il est essentiel de rappeler à quel point le choix des outils appropriés pour accéder à des populations marginalisées ou tout simplement qui ne parlent pas la même langue est déterminant.

Lors des missions sur le terrain, l'équipe d'évaluation sera amenée à déterminer ce qu'il faut observer dans un court laps de temps (contrairement au chercheur d'une enquête ethnographique qui aura le bénéfice d'un temps long d'immersion sur le terrain) : le contexte général de la collecte (à qui et à quoi l'équipe a accès ou pas), des sources historiques (par exemple, consulter le bulletin d'une école sur plusieurs années pour retracer les activités proposées par les services éducatifs de la communauté), de la documentation, des individus, les moments d'interaction (planifiés ou spontanés), la communication non verbale et ses significations, les nuances de langage et les sens donnés à un même mot (qui peut différer de celui de l'équipe d'évaluation même lorsqu'il n'y a pas de différence de langue), des activités ou événements, l'absence d'événements, et soi-même dans une perspective réflexive (Patton, 2014). L'engagement prolongé sur le terrain, la diversification des sources de données (pas seulement les personnes autorisées dans les communautés, et pas seulement les personnes, mais peut-être aussi les lieux eux-mêmes, par exemple l'état d'achèvement d'une école, les archives et la documentation de l'intervention, de l'organisation ou du service évalué), la prise de notes réflexives (dans un journal de bord), l'écriture de descriptions denses et détaillées de ce qui est observé chaque jour, sans oublier les possibilités de prises de photos, dessins, vidéos (avec consentement), etc. sont autant de techniques et procédés qui sont mis en œuvre dans une collecte de données qualitatives pour s'assurer du respect des critères de rigueur scientifique des données qualitatives (Patton, 2014; Fortin et Gagnon, 2016).

L'un des éléments critiques de la présence sur le terrain est de savoir comment se comporter. De la même manière que les personnes envoyées pour la collecte « observent » ce qui se passe, elles sont elles-mêmes observées (quel que soit l'outil mis en œuvre) par les populations auprès desquelles elles interagissent. Au-delà des mots et des présentations des équipes sur le terrain, c'est leur action qui comptera pour la population. Enfin, l'importance que les personnes auprès de qui l'information est collectée accorderont à la collecte dépend de l'intérêt qu'elles y trouveront (sentiment d'importance, sentiment que leur avis compte, plaisir à répondre, ou autres formes de compensation). Patton (2014) appelle ce dernier élément le principe de réciprocité. Lors de l'étape de routinisation de la collecte sur le terrain, il explique également l'importance de gérer un certain équilibre entre la volonté de s'immerger davantage dans l'observation et le besoin de garder une certaine distance avec ce qui est observé. Patton indique que nul n'est automatiquement immunisé contre la tentation de s'impliquer dans la « politique » de l'intervention ou encore contre la possibilité de se retrouver pris en étau dans certains rapports de pouvoir (où la présence de l'équipe sur le terrain est utilisée, parfois à leur insu, à des fins spécifiques par certains membres de la communauté). C'est pour cela que la diversification des sources et la capture de différentes perspectives sont essentielles. D'autres stratégies peuvent à l'inverse impliquer pour l'équipe sur le terrain de s'aligner complètement (sans y adhérer) sur les vues d'un groupe par rapport à un autre afin de bien capter les enjeux, par exemple s'aligner complètement sur les perspectives de participants à des activités de formation (et donc ne pas côtoyer l'équipe du projet présente sur le terrain) (Patton, 2014).

Les enjeux éthiques de la collecte de données devront faire l'objet d'un soin particulier, y compris lors d'une collecte sur Internet (cf. encadré 7).

### Encadré 7 – Les enjeux éthiques des collectes et Internet (Patton, 2014, p. 343)

Patton (2014) met particulièrement en garde ses lecteurs sur les enjeux éthiques des collectes de données sur Internet à partir des recherches de Eysenbach et Till (2001) et Berry (2004) :

- **Caractère intrusif de la collecte (*Intrusiveness*).** Dans quelles circonstances et à quelles fins, s'il en existe, une collecte clandestine (c'est-à-dire sans que les individus sachent qu'elle est en cours) d'informations par des interactions du chercheur/évaluateurs et évaluatrices avec les participants d'une communauté active est-elle considérée éthique ?
- **Respect de la vie privée.** Quelles sont les attentes, perçues ou exprimées, d'un groupe virtuel concernant le respect de la vie privée ? Les propos échangés dans le groupe sont-ils perçus comme privés ? Quelles sont les règles d'adhésion au groupe ?
- **Vulnérabilité et préjudices.** Un certain nombre de groupes virtuels peuvent être affectés à l'appui à des populations vulnérables (victimes de violences sexuelles, d'abus, survivants de maladie, VIH, etc.). Des règles très spécifiques d'encadrement des collectes de données auprès de ces populations, pour éviter tout préjudice et ne pas « faire plus de mal que de bien », ont été réfléchies et mises en pratique, y compris par l'instrument réglementaire, dans la vie réelle. L'évaluateur et évaluatrice doit donc réfléchir au degré de vulnérabilité sur Internet.
- **Consentement.** Le consentement pour participer à l'étude est-il requis ? Comment sera-t-il obtenu ?
- **Confidentialité.** Comment sera protégée l'identité des participants, particulièrement en ce qui concerne les citations dans les supports de présentation des résultats, en sachant que tous les verbatim disponibles sur Internet peuvent être généralement retrouvés ?
- **Propriété intellectuelle.** Dans quelle mesure les participants souhaitent-ils être en contrôle de ce qu'ils disent et de la façon dont leurs propos seront rapportés ? Une pratique d'entrevue, par exemple, est de proposer le verbatim au participant pour confirmer qu'il est en accord avec ce qui est rapporté. Quelle sera la manière de procéder sur Internet ?

Par ailleurs et selon les contextes, il peut s'avérer utile de prendre des intermédiaires choisis au sein de la communauté (à ne pas confondre avec les personnes qui introduisent ou rendent légitimes la présence des équipes sur le terrain, mentionnées dans la section précédente). Ces personnes-ressources constituent un trait d'union entre les populations et les équipes qui vont sur le terrain en tant qu'*informateurs*. Elles peuvent dresser un portrait global du contexte dans lequel va se dérouler la collecte des données et répondre à la fois aux questions de l'équipe sur le terrain (que ce soit directement

l'équipe d'évaluation ou l'équipe d'enquête) et à celles des répondants. Elles sont particulièrement pertinentes dans le cas des missions sur le terrain où des outils qualitatifs de collecte sont plutôt privilégiés, car elles transmettent des informations clés aux personnes sur le terrain, par exemple les moments propices au partage de l'information (fêtes, activité spécifique clé de la journée, événements, etc.), les gestes qui inspirent confiance, ou encore les choses à ne pas faire ou à ne pas dire, etc. Cela dit, le choix de ces personnes-ressources doit être mûrement réfléchi pour éviter de créer l'effet inverse, c'est-à-dire des antagonismes ou de la suspicion (Patton, 2014).

L'ensemble des éléments présentés jusque-là représente une courte synthèse d'un chapitre clé de Patton sur le travail de terrain dans le cadre des stratégies qualitatives de collecte et d'analyse de données (Patton, 2014). Cela dit, voici un exemple concret d'envoi d'une équipe d'enquête, plutôt pour des stratégies quantitatives, basé sur l'expérience d'un des auteurs qui travaille avec l'organisation non gouvernementale (ONG) Innovation Environnement Développement (IED) Afrique<sup>5</sup>.

À l'issue de leur formation, l'équipe d'évaluation répartit l'équipe d'enquête dans les différentes zones ciblées selon des critères spécifiques à l'intervention (distance, taille de la population, temps accordé à l'évaluation, etc.). Les superviseurs sont chargés de prendre contact avec les responsables des villages et avec les autorités compétentes pour partager avec eux toutes les informations pertinentes et les attentes vis-à-vis des participants à l'enquête. Avec l'aide des personnes-ressources, les équipes procèdent à la reconnaissance de leurs zones d'intervention, le choix des ménages et leurs dispositions dans le village, sans oublier la prise en compte des mesures sécuritaires spécifiques, le cas échéant. Chaque enquêteur ou enquêtrice procède à son tour à la reconnaissance de sa zone d'enquête. Dans le village, le travail démarre par une session de sensibilisation avec le chef de village, notamment pour échanger sur la répartition géographique des ménages dans le village, et plus précisément ceux qui font partie de l'échantillon. À ce stade, il est souvent utile d'intégrer les autorités coutumières afin d'éviter de potentiels conflits d'intérêts avec le chef de village. À cette occasion, les enquêteurs et enquêtrices présenteront à nouveau les objectifs de la mission tout en rappelant le soutien attendu des autorités. Le programme de travail des enquêteurs et enquêtrices sera également communiqué en sollicitant tout appui nécessaire pour faciliter l'adhésion des communautés. Une fois les équipes déployées, les superviseurs vérifient la procédure d'échantillonnage utilisée par les équipes avant que les entretiens aient lieu. Au sein du foyer, le succès de l'entretien dépendra du comportement et des attitudes de l'enquêtrice ou de l'enquêteur. Comme précisé plus haut, l'impression que fera l'enquêteur ou l'enquêtrice à son vis-à-vis est déterminante pour obtenir la coopération de ce dernier. L'aspect vestimentaire, le choix des mots, le commencement de l'entretien selon les salutations usuelles, la simplicité et la concision des propos en évitant le jargon, sont autant d'éléments qui permettront de mettre à l'aise l'interlocuteur. Ensuite, l'exposé des motifs de l'enquête constituera un gage d'adhésion des répondants ainsi que l'engagement à la confidentialité, surtout si l'entretien est enregistré. On peut également rappeler qu'il n'y a aucune obligation de répondre aux questions et que l'entretien peut prendre fin à tout moment. Pendant l'entretien, IED Afrique donne une série de consignes : réaliser l'entretien, autant que possible, lorsque le répondant est seul, suivre l'ordre établi, poser la totalité des questions, ne pas précipiter les entretiens, écouter le répondant avec attention, demeurer neutre et impartial tout au long de l'entretien, ne pas suggérer de réponse au répondant, prendre le

---

5. IED Afrique, créée en 2005, est une organisation indépendante, à but non lucratif, basée au Sénégal, qui capitalise une quinzaine d'années d'expérience en Afrique de l'Ouest francophone. Elle intervient sur les problématiques de développement durable et de citoyenneté en Afrique en mettant l'accent sur les innovations méthodologiques et participatives. <http://www.iedafrique.org/>



temps de mettre fin à la conversation de manière courtoise en répondant aux questions du répondant et enfin, ne pas faire de promesses. Il est exigé que les enquêteurs et enquêtrices agissent avec professionnalisme, ce qui implique de coopérer avec les autres membres de l'équipe d'enquête, d'effectuer un nombre spécifique d'entretiens, de prendre le temps nécessaire pour remplir les questionnaires, de réviser les questionnaires remplis à la fin de la journée, de retourner auprès des ménages selon les instructions reçues, de participer à des séances de rétroaction avec les superviseurs et de les informer de l'état d'avancement des travaux.

Chaque exercice de collecte de données doit prévoir un mécanisme de rétroaction utile au suivi et au bilan des activités menées sur le terrain. En cours de collecte, la rétroaction vise surtout à aider les enquêteurs à rendre compte de manière continue des défis rencontrés lorsque les manuels de procédures se révèlent insuffisants face à certaines situations (conflits, difficulté à remplacer certaines unités, épuisement du matériel, etc.). Selon la gravité ou la difficulté de la problématique, la rétroaction pourra se faire auprès du superviseur ou auprès de l'équipe responsable de l'évaluation. Dans tous les cas, il est conseillé de documenter l'ensemble des situations rapportées. Plus spécifiquement, les superviseurs doivent examiner chaque journée avec l'ensemble des enquêteurs et enquêtrices sous leur responsabilité pour discuter des difficultés rencontrées et apporter des mesures correctives. Les superviseurs sont plus particulièrement responsables du contrôle de la présence effective des enquêtrices et enquêteurs sur le terrain, et de l'assurance que ces derniers fassent le travail demandé selon les orientations données, de la localisation géographique des enquêteurs et enquêtrices en temps réel, du contrôle qualité pendant les entretiens (pour apporter des corrections sur la technique d'entretien de l'enquêteur ou l'enquêtrice), de la réception à la fin de la collecte de l'ensemble des dossiers pour vérifier l'exhaustivité et la qualité du travail accompli, de l'élaboration d'un rapport de la collecte dans sa zone de supervision, incluant les difficultés rencontrées et leurs résolutions.

Indépendamment des outils et des techniques déployés, il est important pour l'équipe sur le terrain de prendre le temps de clore la collecte, c'est-à-dire de prendre des dispositions par rapport aux intermédiaires ou personnes-ressources clés, par rapport aux représentants des personnes rencontrées, etc. Dans le cadre d'une stratégie de collecte qualitative, bien que la collecte en elle-même ait lieu plutôt selon un processus itératif entre la découverte et la validation des informations collectées (Guba, 1978), à la fin de la période sur le terrain, l'équipe procédera à une vérification des données collectées plutôt qu'à faire émerger de nouvelles informations. Un mécanisme de rétroaction pour vérifier l'information collectée peut être prévu, par exemple sous la forme de la transmission d'une description de l'information collectée auprès des populations ou des équipes de projet dont est issue l'information. Les réactions à ces éléments peuvent alors être incluses dans l'analyse. Attention, l'évaluatrice ou l'évaluateur doit demeurer prudent à cet égard et faire la différence entre la validation de la synthèse de l'information collectée par la personne dont est issue l'information des constats de l'évaluation.

Finalement, à l'issue de la collecte de données, une rencontre regroupant l'ensemble des personnes qui ont réalisé la collecte d'information (enquêteurs, facilitateurs, superviseurs) et l'équipe d'évaluation peut être organisée afin que chacune des personnes impliquées fasse part des défis rencontrés, de certaines anecdotes et des bons coups réalisés.

## Conclusion

Ce chapitre porte une attention particulière sur des pratiques concrètes d'entrées sur le terrain assez classiques (orientées sur l'administration de questionnaires en personne et l'observation), en particulier lorsque celles-ci sont physique, même si certains éléments abordés font référence aux possibilités numériques ou à distance de collecte de données. Quel que soit le choix de terrain, qu'il soit restreint ou élargi, la collecte de données implique une diversité de sources possibles, et nécessite une préparation attentive aux possibilités offertes par le contexte de l'évaluation, qui n'est pas celui de la recherche. Or, il s'agit du moment de l'évaluation le plus concret pour les populations cibles, celui où elles sont directement en contact avec le mot « évaluation ». Il est donc dommage que ce moment soit perçu comme une simple étape de mise en œuvre de la démarche, souvent déléguée, alors que cette étape est non seulement essentielle pour la réussite de l'évaluation, mais elle engage également la réputation du domaine auprès des populations.

## Messages clés

1. L'entrée sur le terrain nécessite une préparation adéquate qui dépend du contexte de l'évaluation, notamment du client principal (et/ou du comité d'évaluation), de la zone ciblée par la collecte (accessibilité), des us et coutumes associés aux populations de cette zone, de l'outil choisi en fonction des populations à rencontrer, des législations et procédures administratives, des calendriers des populations à rencontrer, des langues locales, des ressources disponibles et des délais de l'évaluation.
2. La nature et la manière de communiquer l'information concernant l'évaluation est de nature stratégique et ne peut être décidée par la seule équipe d'évaluation. Elle doit faire l'objet d'une réflexion approfondie.
3. La présence sur le terrain n'est pas neutre. Les équipes présentes sur le terrain sont tout autant observées qu'elles observent à leur tour. Les mauvaises expériences de collecte de données contribuent à la méfiance des populations. Le principe de réciprocité, l'éthique de la collecte, la pratique de la rétroaction, etc. sont autant d'éléments qui contribuent à rendre l'expérience positive pour les personnes impliquées.
4. L'évaluateur ou l'évaluatrice doit bien maîtriser les techniques et procédés propres à chaque nature d'outil, qualitatif ou quantitatif, pour respecter les critères de rigueur scientifique propres à chacun.
5. La délégation de la collecte de données à des équipes d'enquête nécessite un recrutement et une préparation adéquate de ces équipes.

## Références bibliographiques

- Affaires mondiales Canada. (2016). *La gestion axée sur les résultats appliquée aux programmes d'aide internationale : un guide pratique* (2<sup>e</sup> édition). Centre d'excellence de la gestion axée sur les résultats. [https://www.international.gc.ca/world-monde/assets/pdfs/funding-financement/results\\_based\\_management-gestion\\_axee\\_resultats-guide-fr.pdf](https://www.international.gc.ca/world-monde/assets/pdfs/funding-financement/results_based_management-gestion_axee_resultats-guide-fr.pdf)
- Bensa, A. (2006). *La fin de l'exotisme*. Éditions Anacharsis.
- Berry, D. M. (2004). Internet research : Privacy, ethics and alienation. *Internet Research*, 14(4), 323-332. <https://doi.org/10.1108/10662240410555333>
- Blanc, G. (2012). Une pratique sans questionnement : le terrain en histoire. *Hypothèses*, 1(15), 15-25. <https://doi.org/10.3917/hyp.111.0015>
- Conseil international de la langue française. (1976). *Terrain*. Office québécois de la langue française. [https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id\\_Fiche=17575272](https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=17575272)
- Dufour, S., Fortin, D. et Hamel, J. (1991). *L'enquête de terrain en sciences sociales : l'approche monographique et les méthodes qualitatives*. Éditions Saint-Martin.
- Eysenbach, G. et Till, J. E. (2001). Ethical issues in qualitative research on internet communities. *British Medical Journal*. <https://doi.org/10.1136/bmj.323.7321.1103>
- Fortin, M.-F. et Gagnon, J. (2016). *Fondements et étapes du processus de recherche : Méthodes quantitatives et qualitatives* (3<sup>e</sup> édition). Chenelière Éducation.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, R. B. et Vermeersch, M. J. C. (2011). *L'évaluation d'impact en pratique*. Banque mondiale.
- Given, M. (dir.) (2008). *The SAGE Encyclopedia of Qualitative Research Methods*. SAGE Publications.
- Guba, E. G. (1978). *CSE Monograph Series in Evaluation No 8. Toward a methodology of naturalistic inquiry in educational evaluation*. University of California. Center for the Study of Evaluation. UCLA Graduate School of Education. [http://cresst.org/wp-content/uploads/cse\\_monograph08.pdf](http://cresst.org/wp-content/uploads/cse_monograph08.pdf)
- Jacob, S. et Ouvrard, L. (2009, automne). L'évaluation participative. Avantages et difficultés d'une pratique innovante. *Cahiers de la performance et de l'évaluation*, n° 1.
- Laurent, J. (2009). L'anthropologie à l'épreuve de la décolonisation de la recherche dans les études autochtones. *Anthropologie et Sociétés*, 3(23), 179-196. <https://doi.org/10.7202/029723ar>
- Marchive, A. (2012). Introduction. Les pratiques de l'enquête ethnographique. *Les sciences de l'éducation*, 45(4). <https://doi.org/10.3917/lstdle.454.0007>
- Patton, M. Q. (2014). *Qualitative Research and Evaluation Methods* (4<sup>e</sup> édition). SAGE Publications.
- Patton, M. Q. et Campbell-Patton, C. E. (2022). *Utilization-Focused Evaluation* (5<sup>e</sup> édition). SAGE Publications.
- Salün, M. (2012). L'ethnographie à l'heure des « bonnes pratiques ». Retour sur une expérience tahitienne. *Les sciences de l'éducation*, 45(4), 95-112. <https://doi.org/10.3917/lstdle.454.0095>
- Statistique Canada. (2010). *Méthodes et pratiques d'enquête*. [https://www150.statcan.gc.ca/n1/fr/pub/12-587-x/12-587-x2003001-fra.pdf?st=qF\\_OYxjQ](https://www150.statcan.gc.ca/n1/fr/pub/12-587-x/12-587-x2003001-fra.pdf?st=qF_OYxjQ)
- Weber, F. (2015). *Brève histoire de l'anthropologie*. Flammarion.



# L'analyse des données pour répondre aux questions d'évaluation



*Miché Ouédraogo*

*Isabelle Brunet*

*Wend-Yam Saturnin Zoetyande*

*Maxime Desfossés*

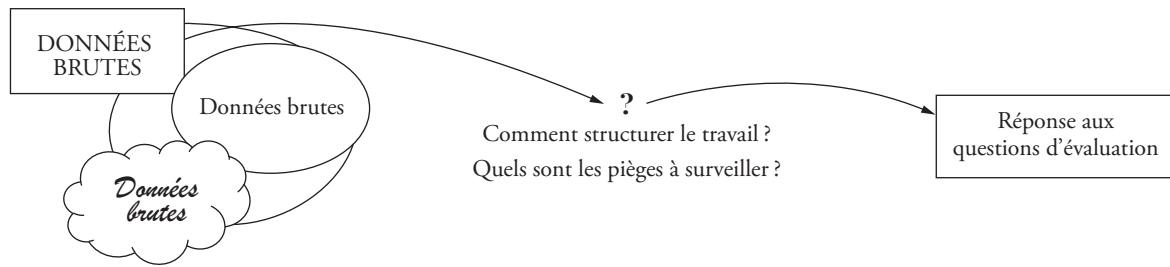
*Vénétia Sauvain*

## Introduction

Se présentant comme UNE phase de l'évaluation, on pourrait croire à tort que l'analyse se réalise à travers UNE tâche qui consisterait par exemple à appliquer des formules pour produire des statistiques, à concevoir des tableaux ou à faire une synthèse des points de vue des parties prenantes. Comme l'explique Fortin (2016) lorsqu'il aborde la démarche qualitative, la phase analytique se conçoit et s'exécute plutôt à travers **un parcours** aux multiples tâches dans un mouvement itératif. C'est la rigueur de la démarche analytique qui distingue l'analyse évaluative d'une simple démarche de questionnement ou de remise en question d'une initiative. Ce chapitre aborde la structuration de la phase analytique, les principaux procédés d'analyse qualitative et quantitative utilisés en évaluation, ainsi que quelques idées pour l'intégration de ces analyses dans une perspective de stratégie mixte d'analyse des données pour répondre aux questions d'évaluation.

## La phase analytique : un parcours en soi

Tel qu'illustré plus bas, ce parcours de la phase analytique a pour point de départ des données brutes dépourvues de réelle signification et pour point d'arrivée la réponse aux questions d'évaluation, c'est-à-dire l'éclairage dont ont besoin les responsables du programme. Entre ces deux points s'exécutent des travaux et apparaissent certains enjeux qu'il importe d'explorer ici pour y voir clair pour préserver la crédibilité et l'utilité de l'évaluation en elle-même.

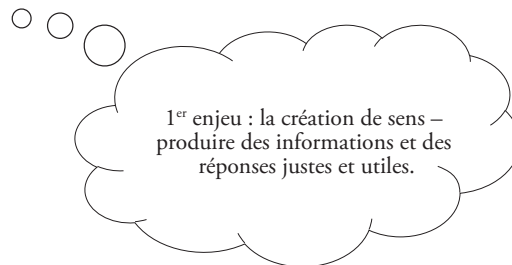


**Figure 1**

Le parcours analytique : un travail de transformation.

Dans l'ordre simple des choses, on pourrait dire qu'après avoir collecté les données : 1) l'évaluateur ou l'évaluatrice les traite; 2) les analyse; 3) les interprète. Ceux qui se sont adonnés à l'exercice conviendront que dans la pratique, ces travaux se réalisent souvent simultanément, allant même parfois jusqu'à se chevaucher et se confondre les uns aux autres. Et c'est encore plus évident lorsque l'on aborde des questions d'évaluation descriptives ou explicatives qui font appel à l'induction.

Plus précisément, on pourrait aussi dire que le traitement des données consiste essentiellement en un travail d'extraction, de conversion, d'organisation, de codification, de calcul (en appliquant des formules) et de montage de tableaux et de graphiques. En effectuant ce travail de traitement, l'intention principale est de transformer les données brutes en informations utiles et ainsi accéder à une représentation juste et complète de l'objet d'évaluation. On dira aussi que sur la base de ces informations, il est maintenant possible de procéder à l'analyse proprement dite, qui consisterait alors à s'appropriier les informations. La façon de s'appropriier les informations sera évidemment tributaire de la méthode retenue lors de la planification de l'évaluation – cette méthode découlant elle-même du type de questions (descriptives, explicatives, normatives) – c'est-à-dire de la position épistémologique de départ (Cherkaoui et Haouata, 2017). En réalisant ce travail d'analyse, on cherche essentiellement à faire des constats – c'est-à-dire constater ce qu'il en est de l'objet d'évaluation. Enfin, l'interprétation consisterait à donner du sens précis aux principaux constats, on cherche ici à se positionner face à ce que l'on observe, à se faire une opinion et ultimement à produire une réponse juste aux questions d'évaluation.



**Tableau 1** — Étapes de la phase analytique

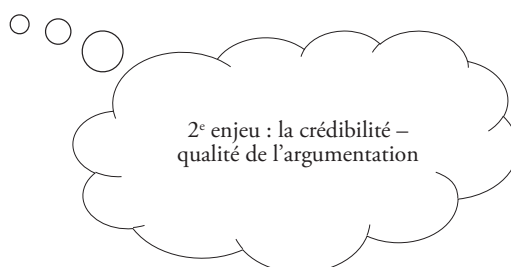
	Le travail à réaliser	Le résultat
Traitement	<b>Calculer</b> Moyenne, %, somme – appliquer la formule, additionner le nombre d’hommes, de femmes.	Des informations.
Analyse	<b>Comparer</b> – référence à la norme (qui peut être un objectif) On constate que les hommes ... et que les femmes.... alors que les jeunes...	Des constats.
Interprétation	<b>Répondre</b> Les résultats observés démontrent que les interventions auprès des jeunes permettent...	Une réponse. Une opinion, une position.

### Astuces pour structurer la phase analytique

Afin de s’assurer de l’utilité des réponses et de la capacité de les argumenter avec crédibilité, il convient de préparer et de réaliser la phase analytique en :

1. Tentant de distinguer les étapes de traitement (produire des informations), d’analyse (produire des constats), et d’interprétation des données (produire des réponses);
2. Assurant un parcours analytique distinct et spécifique à chacune des questions d’évaluation. Bien que différentes questions fassent partie du même protocole (devis), chacune d’elles doit trouver réponse selon une stratégie qui lui est propre;
3. Prévoyant des critères selon lesquels l’évaluateur ou l’évaluatrice retiendra ou ne retiendra pas certaines informations pour répondre à la question d’évaluation (plus une justification de ces critères).

Cette distinction entre le traitement, l’analyse et l’interprétation devient toutefois plus difficile à faire lorsque l’on compose avec des données qualitatives où le travail de traitement et celui d’analyse se conjuguent. Un des pièges fréquents de cette phase est celui de répondre aux questions d’évaluation trop hâtivement sans que les constats aient été faits, discutés et validés. Dans un tel cas, l’évaluateur ou l’évaluatrice compromet sérieusement l’acceptabilité de la réponse qu’il obtient et sa capacité à présenter une argumentation crédible pouvant la soutenir (Houle et Hurteau, 2012, p.187). L’évaluateur ou l’évaluatrice doit absolument éviter ce piège, car la phase d’analyse, comme l’ensemble de la démarche d’évaluation, ne doit pas uniquement donner lieu à une réponse, elle doit également pouvoir présenter la façon dont cette réponse a été produite.

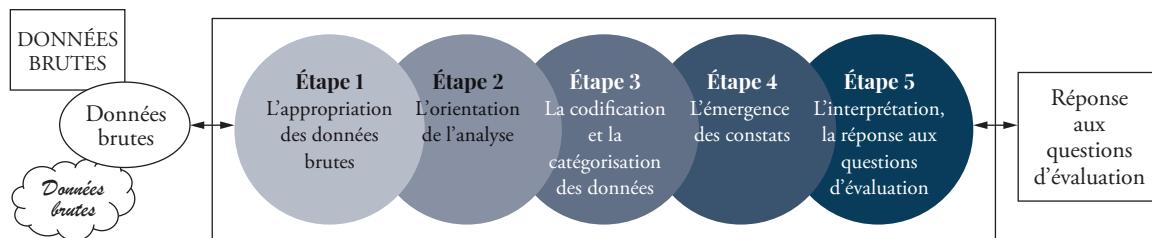




## L'analyse thématique des données qualitatives pour répondre aux questions d'évaluation : de la codification à l'interprétation

Lorsque l'on entreprend une évaluation avec des données qualitatives, c'est souvent la phase analytique qui s'avère la plus laborieuse. Cette complexité s'explique en partie par le fait que l'analyse des données qualitatives « [...] passe par la réduction d'une importante quantité de matériel dont la nature est souvent diversifiée » (Turcotte *et al.*, 2009, p. 207). L'évaluateur ou l'évaluatrice structurera alors son parcours analytique sachant que l'unique ou la meilleure façon de procéder n'existe pas. Tout comme dans le cas d'une démarche quantitative ou mixte, ses choix doivent se faire et se préciser principalement en fonction de la question d'évaluation et de l'utilisation prévue de la réponse. Le temps et les moyens dont il dispose sont aussi d'importants éléments à prendre en compte afin d'assurer la faisabilité des travaux d'analyse.

Parmi la multitude d'options, on retient ici la séquence d'analyse des données qualitatives que Taylor-Powell et Renner (2003) proposent et qui s'apparente à celle présentée plus tôt dans ce chapitre. C'est-à-dire qu'il s'agit toujours essentiellement de transformer les données brutes en informations utiles; d'en tirer des constats et de produire une réponse à la question d'évaluation.

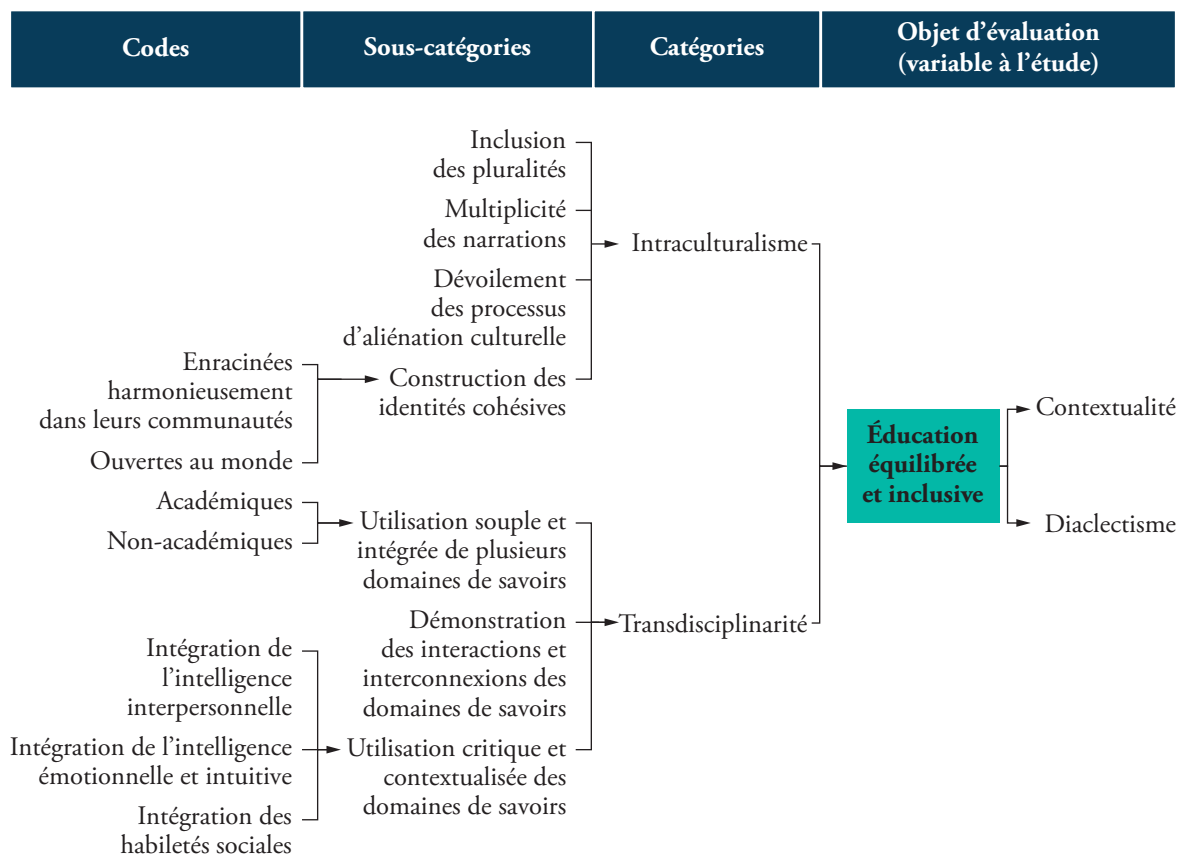


**Figure 2**

Schéma synthétique du processus d'analyse qualitative.

Voici donc ce qu'il en est de ce parcours, lorsqu'adapté aux spécificités de la démarche qualitative sous l'angle de l'analyse thématique où les données se présenteraient sous forme de mots, de phrases, d'extraits de texte issus de témoignages, d'entretiens, de discussions de groupes, de journaux, de notes d'observations, de rapports, d'histoires ou d'études de cas.

Cette analyse vise à reprendre les thèmes clés en lien avec les variables étudiées dans la matrice d'évaluation pour élaborer des synthèses à partir des données qualitatives collectées (mots, phrases, etc.). Comportant généralement un niveau de complexité élevé, les objets d'évaluation (variables étudiées) se déclinent en plusieurs codes, lesquels se regroupent en catégories et en sous-catégories. La carte heuristique, aussi appelée *Mindmap*, est fort utile pour schématiser cette déclinaison. À titre d'exemple, en supposant que la démarche d'évaluation consiste à *comprendre comment se concrétise et s'expérimente l'éducation équilibrée et inclusive dans le cadre d'interventions qui visent ce résultat*, l'analyse thématique pourra se réaliser en fonction du cadre conceptuel existant sur le sujet dans la littérature.



**Figure 3**

Carte heuristique des thématiques associées à l'éducation équilibrée et inclusive.

Tel que présenté sur la carte, l'objet d'évaluation « éducation équilibrée et inclusive » (c'est-à-dire la variable à l'étude dans la matrice, voir le chapitre « La matrice d'évaluation ») se définit à travers une multitude de codes, catégories et sous-catégories. Dans le cas où l'évaluateur ou l'évaluatrice opérerait pour une approche déductive (c'est-à-dire en utilisant des codes prédéfinis), une telle carte conceptuelle servira de grille pour procéder à l'analyse thématique des données qualitatives dont il dispose.

**Étape 1 – L'appropriation des données brutes.** Considérant que la qualité de l'analyse repose d'abord sur la compréhension que l'évaluateur ou l'évaluatrice a des données, la première étape consiste ainsi à s'approprier les écrits, c'est-à-dire les lire, les relire plusieurs fois et par la même occasion en estimer la valeur. Pour cette étape, il est essentiel d'avoir une approche compréhensive par rapport aux participants, en ne faisant pas une interprétation subjective des propos rapportés, tout en gardant une attitude exempte de biais ou de préjugés (Andreani et Conchon, 2005). L'évaluateur ou l'évaluatrice pourra déjà ici identifier certaines limites du matériel dont il dispose et les documenter. Une des principales limites méthodologiques rencontrées en évaluation est l'absence de verbatim exact des entrevues et des groupes de discussion menés. Souvent par manque de temps, l'équipe d'évaluation aura procédé à des verbatim approximatifs ou des comptes rendus synthétiques, ce qui limite les analyses thématiques quantitatives (en effet, l'analyse des occurrences d'un thème est alors impossible à déterminer).

**Étape 2 – L'orientation de l'analyse.** À cette étape, le travail de l'évaluateur ou l'évaluatrice consiste à faire un choix sur la façon dont l'analyse sera orientée, structurée. Deux options s'offrent à lui : orienter l'analyse en fonction du contenu (questions, objets, périodes ou événements) ou en fonction des répondants (cas, individus, groupes). Le choix se fera en considérant attentivement les besoins d'information ayant motivé l'évaluation, les questions d'évaluation elles-mêmes et l'utilisation que l'on prévoit faire des réponses.

**Étape 3 – La codification et la catégorisation des données.** Le travail de codification consiste à associer les données à des thèmes (codes). C'est ainsi que l'évaluateur ou l'évaluatrice donne un sens aux mots, aux phrases, aux extraits issus des textes, mais également aux images ou aux données visuelles; bref, à l'ensemble des données qualitatives dont il dispose. Selon la nature de l'information à produire, ces codes peuvent renvoyer à des idées, des concepts, des comportements, des interactions, des incidents. L'évaluateur ou l'évaluatrice peut procéder par **déduction** en utilisant des thèmes prédéfinis liés au modèle théorique retenu (comme dans l'exemple de l'éducation équilibrée et inclusive) ou procéder par **induction** en laissant les thèmes émerger du matériel dont il dispose. Il est également possible que des thèmes non prévus initialement émergent des données collectées, et il est essentiel de les traiter si ces données permettent d'apporter un éclairage nouveau par rapport à une ou des variables de la matrice.

Dans certains cas, l'évaluateur ou l'évaluatrice conjugue les deux façons, en débutant par exemple avec des thèmes préétablis, tout en laissant une ouverture pour l'émergence d'autres thèmes qu'il n'aurait pas choisis au départ. La catégorisation consisterait alors à mettre en relation les thèmes identifiés et à les regrouper selon ce qu'ils ont en commun, ce qu'ils présentent comme affinités. Cette opération délicate de codification/catégorisation constitue le cœur même de la phase analytique qualitative. Elle fait appel à un haut niveau de rigueur de la part de l'évaluateur ou de l'évaluatrice pour s'assurer de l'adéquation des thèmes/catégories ainsi que de leur logique et de leur exhaustivité.

### La qualité de la codification

L'analyse des données se doit d'être rigoureuse pour garantir la validité des résultats (Nations Unies, 2004). Comme premier critère de contrôle, on retrouve l'exhaustivité : a-t-on réussi à classer tous les codes dans les thèmes créés? Ensuite, il importe d'évaluer l'objectivité des codes : sont-ils suffisamment clairs et compréhensibles de la même façon par les divers codeurs? Les données doivent être codées de façon standardisée et consistante pour s'assurer de la validité des résultats. De plus, en s'assurant de l'objectivité de la codification, il sera plus aisé de diviser le travail sans inquiétudes quant à la conformité de la codification entre les codeurs (Delve, 2021). Finalement, il faut également s'assurer de la pertinence de la codification : est-ce que les codes sont utiles pour répondre aux questions d'évaluation? En ce sens, les codes doivent être en harmonie avec le cadre conceptuel préalablement établi.

**Étape 4 – L'émergence des constats.** En associant les données aux thèmes et en les catégorisant, l'évaluateur ou l'évaluatrice voit apparaître certains phénomènes – il peut s'agir de la récurrence de certains thèmes, de relations intracatégories et de relations intercatégories. Pour le chercheur, il s'agit donc d'être attentif aux associations et dissemblances possibles entre les catégories et entre les thèmes. À ce moment de l'analyse, certains thèmes sont appelés à disparaître, à fusionner ou encore à

apparaître (Paillé et Mucchielli, 2021). C'est sur cette base que l'évaluateur ou l'évaluatrice sera à même de constater ce qui en est de l'objet de l'évaluation. La quantification des occurrences peut s'avérer un exercice utile pour faire valoir l'importance de certains thèmes ou catégories par rapport à d'autres, à condition de disposer des verbatim exacts. Les logiciels de traitements de données qualitatives tels que NVivo et QDA Miner sont particulièrement utiles à ce niveau puisqu'ils permettent de générer des statistiques et faire de la modélisation. Néanmoins, il convient de préciser que ces logiciels peuvent être coûteux. Des alternatives telles que les tableurs (chiffriers) restent envisageables. De plus, ils ne se substituent pas au travail d'analyse lui-même.

**Étape 5 – L'interprétation, la réponse aux questions d'évaluation.** En s'élevant au-dessus de l'ensemble des constats, l'évaluateur ou l'évaluatrice doit maintenant se positionner. C'est cette position qui lui permettra de produire sa réponse aux questions d'évaluation, de la documenter et de l'argumenter.

À cette étape, il est impératif que l'évaluateur ou l'évaluatrice structure son travail de manière à assurer la cohérence entre les constats dont il dispose, les informations qui y sont liées et la question d'évaluation. Toutefois, souvent trop nombreux, les résultats issus de l'analyse doivent faire l'objet d'un choix. Quelles informations retenir, lesquelles ne pas retenir? Comment choisir parmi les nombreux constats? Pour faire ces choix, il s'avère donc important que l'évaluateur ou l'évaluatrice définisse des critères selon lesquels il retient ou non certains constats et certaines informations (Houle et Hurteau, 2012, p. 207). Il s'agit d'un travail crucial, considérant que la crédibilité de l'analyse et par le fait même, celle des réponses, sont en jeu. Pour la crédibilité de l'étude, il est important que le chercheur aborde tous les résultats pertinents en lien avec la question ou les questions d'évaluation et le cadre conceptuel. Les résultats qui n'étaient pas attendus ou qui ne correspondent pas aux explications données du phénomène dans la littérature ont leur pertinence et peuvent s'avérer être des pistes de réflexion vers d'autres enjeux (Nowell *et al.*, 2017). Il s'agit donc pour l'évaluateur ou l'évaluatrice d'examiner les explications rivales et de procéder à une comparaison avec d'autres données et d'autres sources pour formuler une réponse à la question d'évaluation. De plus, les ambiguïtés dans les résultats peuvent être signe de la complexité du phénomène étudié, sans rendre les résultats de l'évaluation moins valides : sa validité repose davantage sur sa méthodologie et son explication en profondeur (Antin *et al.*, 2014).

Enfin, afin de faciliter l'acceptabilité des réponses et la compréhension des arguments par les utilisateurs de l'évaluation, il convient de présenter des figures démontrant l'importance de certaines informations ou de présenter des graphiques illustrant les relations intra ou intercatégorie. Pouvant hiérarchiser et modéliser, les logiciels dédiés au traitement et à l'analyse des données qualitatives seront encore une fois particulièrement utiles sur ce plan. Toutefois, ces analyses qualitatives poussées ne sont pas toujours nécessaires et il convient de bien déterminer si elles seront utiles pour répondre aux questions d'évaluation.

### **Favoriser le contrôle qualité des constats issus de données qualitatives (Patton, 2008).**

- Dans quelle mesure mon protocole est-il rigoureux et cohérent pour soutenir les constats?
- Dans quelle mesure les constats permettent-ils de mieux comprendre le phénomène étudié?
- Dans quelle mesure les constats sont-ils cohérents avec d'autres résultats connus?
- Dans quelle mesure les constats ont-ils du sens pour les utilisateurs? Et sont-ils utiles à la prise de décision?

## L'analyse des données quantitatives pour répondre aux questions d'évaluation

Comme les données qualitatives, les données quantitatives doivent être traitées (vérification des « ne sais pas », « ne veux pas répondre » ou « ne s'applique pas » le cas échéant, nettoyage des valeurs aberrantes, traitement des données manquantes, etc.), analysées et interprétées afin d'apporter les réponses aux questions d'évaluation. Il existe essentiellement trois procédés d'analyse des données quantitatives que nous aborderons dans cette section : les analyses univariées, les analyses bivariées et les analyses multivariées. Mais avant d'aborder ces procédés, il convient de distinguer les différentes formes de données quantitatives que l'on peut rencontrer dans le cadre de ces analyses.

### La particularité des données quantitatives : numérique, ordinale, de ratio, etc.

La variable se définit comme « une caractéristique ou une propriété quelconque dont la valeur diffère d'un cas à l'autre » (Fox, 1999). La variable est donc ce qui caractérise un aspect d'un individu. Les données sont exploitées sous forme de variables à partir de logiciels d'analyses de données quantitatives (Excel, SPSS, STATA, R, etc.). Dans tous ces logiciels, les variables sont exprimées sous forme de colonnes, et les individus sous forme de lignes. Les différentes possibilités de réponses prises par ces variables sont appelées modalités. On étudie un échantillon ou une population à partir de différents types de variables : nominales, ordinales, numériques à intervalles et numériques de rapport (présenté dans le tableau ci-dessous). Il est important de comprendre la nature des variables, car elle détermine les analyses statistiques qui peuvent être menées.

**Tableau 2** — Exemples de type de variables

Traitement statistique : Faible		→			Élevé
Types de variables	Nominale	Ordinale	Intervalle/ratio	Numérique de rapport	
Définition	Les modalités de ces variables sont des caractéristiques de l'individu / des catégories qualitatives (genre, nationalité, etc.) <sup>1</sup> .	Les modalités de ces variables peuvent être rangées/ ordonnées les unes par rapport aux autres. Généralement, il existe des échelles de valeurs préétablies.	Variables numériques dont les intervalles entre les unités de mesure sont constants d'un point à un autre, ce qui rend possibles les soustractions et les additions.	Variables numériques dont les modalités sont composées d'un zéro absolu, c'est-à-dire que toutes les opérations mathématiques sont possibles.	
Exemples	Villes d'intervention d'un programme	Appréciation d'activité d'un programme	Date d'une formation	Dépenses de santé des ménages en euros	
Exemples de modalités	<ul style="list-style-type: none"> <li>▪ Ouagadougou</li> <li>▪ Abidjan</li> <li>▪ Cotonou</li> <li>▪ Yaoundé</li> </ul>	<ul style="list-style-type: none"> <li>▪ Très satisfait</li> <li>▪ Assez satisfait</li> <li>▪ Peu satisfait</li> <li>▪ Pas du tout satisfait</li> </ul>	<ul style="list-style-type: none"> <li>▪ 2018</li> <li>▪ 2019</li> <li>▪ 2020</li> <li>▪ 2021</li> </ul>	<ul style="list-style-type: none"> <li>▪ 1500 euros</li> <li>▪ 2 500 euros</li> <li>▪ 750 euros</li> </ul>	

1. On peut également retrouver ces catégories, dénommées « attribut » dans les analyses qualitatives, lorsque l'on veut catégoriser les participants aux entretiens ou groupes de discussion.

Dépendamment du type de variable et de la taille de l'échantillon ou de la population étudiée, des analyses univariées, bivariées ou multivariées peuvent être menées. Les deux premières sont qualifiées d'analyses descriptives, et la dernière d'analyse explicative.

### Les analyses univariées ou descriptives : comprendre les paramètres de tendances centrales et de dispersion d'une variable

Les analyses univariées sont des analyses statistiques que l'on réalise sur une seule variable. Ce type d'analyse nous permet de voir comment est la distribution de la variable au sein de notre échantillon d'étude à partir d'un simple tableau de fréquence, du calcul des paramètres de tendance centrale ou de dispersion.

Un tableau de fréquence permet de voir, par exemple, le nombre de personnes pour chacune des modalités de la variable. Il est donc possible, sur cette base, de calculer en pourcentage la répartition des individus pour chaque modalité. Les paramètres de tendance centrale sont la moyenne, la médiane et le mode (voir le tableau ci-dessous).

**Tableau 3 — Tableau synthèse des analyses descriptives**

Analyse univariée	Définition	Avantages et limites	Niveau de mesure de la variable requis
<b>Paramètre de tendance centrale</b>			
Moyenne	Mesure de tendance centrale que l'on obtient en additionnant tous les scores et en divisant ensuite cette somme par le nombre de scores.	Du fait qu'elle utilise pour son calcul toutes les valeurs prises par la variable, la moyenne arithmétique peut s'avérer utile dans bien des contextes. La moyenne n'a de sens que pour des variables quantitatives. La moyenne arithmétique présente l'inconvénient d'être sensible aux valeurs extrêmes (valeurs aberrantes).	Intervalle/ratio
Médiane (Md)	Correspond à la valeur qui divise en deux parties égales un ensemble ordonné de scores.	Son calcul est facile. Donne une idée satisfaisante de la tendance centrale de la distribution. Tient compte seulement de l'ordre de grandeur des valeurs prises par la variable. Concerne uniquement les variables quantitatives.	Intervalle/ratio
Mode (Mo)	Modalité qui apparaît le plus fréquemment pour une variable donnée. Lorsqu'un seul des scores apparaît clairement comme le plus commun, nous disons de la variable qu'elle est unimodale.	La détermination du mode est aisée (graphiquement). Son intérêt est évident puisqu'il désigne la valeur de la variable qui est la plus observée sur l'échantillon. Le mode n'a de signification véritable que si l'effectif correspondant est nettement supérieur aux effectifs des autres modalités. En évaluation, le mode est plus intéressant lorsqu'il est unique.	Nominal Ordinal Intervalle/ratio

Les paramètres de dispersion sont la variance et l'écart-type. La variance peut être définie comme étant la moyenne des écarts au carré des scores par rapport à la moyenne. Si les scores de la variable se distribuent de façon étendue autour de la moyenne, les écarts seront grands et, ainsi, la variance sera élevée. Dans le cas contraire, les écarts et la variance seront faibles. Donc, une forte variance témoigne d'une grande variation, une faible variance d'une variation plus faible. L'écart-type, quant à lui, est la racine carrée de la variance. Il mesure en moyenne la variation des scores d'une variable autour de la moyenne. Une variance ou un écart-type nul sous-tend qu'il n'existe pas de variation sur les valeurs de la variable dans notre échantillon. Notre échantillon est donc homogène pour la variable considérée.

Pour l'évaluation, ces analyses descriptives permettent d'identifier les données manquantes ou les valeurs extrêmes. Il est important de les corriger sinon elles peuvent fausser les calculs des paramètres de tendance centrale et de dispersion. Ainsi, les analyses univariées correspondent à la première étape de lecture de l'évaluateur ou de l'évaluatrice pour déterminer les possibilités d'analyse plus poussées à mener.

### **Les analyses bivariées : comprendre la relation entre deux variables quantitatives**

On fait de l'analyse bivariée lorsque l'on croise deux variables pour voir comment elles se comportent et pour interpréter les résultats qui en ressortent. Le tableau de contingence présente la combinaison des catégories de deux variables. La dimension du tableau de contingence est donnée par le nombre de modalités des deux variables qui la composent. Les lignes et les colonnes sont constituées par les modalités des variables. Dans un tableau de contingence, on désigne par fréquences jointes ou encore fréquences d'association le nombre d'observations qui se retrouvent à la fois dans deux catégories des variables du tableau. On appelle « marginales » ou « fréquences marginales » les fréquences absolues et relatives correspondant aux catégories de chacune des variables qui entrent dans la constitution du tableau. Les fréquences conditionnelles sont constituées de la distribution obtenue pour une variable lorsque se réalise une condition définie par la catégorie d'une autre variable, ou par la combinaison des catégories d'autres variables. Les fréquences marginales absolues ou relatives sont des fréquences non conditionnelles. Déterminer la significativité statistique d'une mesure d'association revient à déterminer dans quelle mesure la relation entre deux variables dans un échantillon n'est pas due au hasard, mais existe réellement dans la population de laquelle l'échantillon est extrait. Le terme  $H_0$ , nommé hypothèse nulle, est le postulat selon lequel il n'y a pas de relation dans notre population. La probabilité que l'on retrouve grâce au seul hasard une relation à l'intérieur d'un échantillon alors qu'il n'en existe pas dans la population est appelée niveau de signification ( $p$ ). Il varie entre 0 et 1. Par convention, les chercheurs retiennent qu'une probabilité inférieure ou égale à 0,05 est suffisante pour conclure la signification statistique d'une relation. Ainsi si  $p \leq 0,05$ , on rejette  $H_0$ , et si  $p \geq 0,05$ , on accepte  $H_0$ .

Pour réaliser des analyses statistiques plus poussées, il est important de bien comprendre la notion de variables dépendante et indépendante dans la modélisation statistique. Dans une modélisation statistique, la variable résultante, qui est le résultat, ou les conséquences d'autres variables, est appelée variable dépendante ou variable à expliquer. Cependant, les variables qui concourent à un résultat, à une situation, qui permettent de comprendre un comportement, sont appelées variables indépendantes ou explicatives. Ces notions de variables indépendantes, variables dépendantes ne sont pas toujours intrinsèques à la variable, mais dépendent de la construction théorique de la problématique. Une variable peut être une variable dépendante dans une situation donnée et être une variable indépendante dans une autre situation. En évaluation, on tente plus souvent de déterminer si les variables sont indépendantes et de juger de la force de l'association par la différence entre les pourcentages.



## Les analyses multivariées, analyses explicatives : comprendre la relation entre plusieurs variables et la modélisation statistique

Les analyses multivariées sont utilisées afin de comprendre comment les variables agissent les unes sur les autres lorsque l'on met en relation plus de deux variables. Par exemple, lorsque l'on cherche à comprendre quels sont les facteurs qui ont plus ou moins influencé le succès ou l'échec d'un programme, on a recours à une analyse multivariée où l'on tente de modéliser le succès du programme. On parle de régression linéaire multiple lorsque l'on a plusieurs variables indépendantes dans notre modèle et que nous voulons analyser la contribution de l'ensemble de ces variables à la variation de notre variable indépendante. Par exemple, dans le cadre de l'évaluation d'un programme, si on veut comprendre les déterminants de l'augmentation du revenu, alors l'on aura recours à une régression linéaire dans laquelle le revenu sera la variable à expliquer et les facteurs que l'on veut identifier comme variables explicatives. Il s'agit d'une méthode avancée des évaluations d'impact contrefactuel (voir chapitre 13 : *Les méthodes qualitatives et quantitatives pour évaluer les effets*). Finalement, l'évaluateur ou l'évaluatrice a donc à sa disposition une panoplie de procédés d'analyse pour répondre à des questions d'évaluation explicative.

Dans le tableau suivant, vous retrouverez des exemples, des définitions et des interprétations de quelques mesures d'associations utilisées en évaluation.

**Tableau 4 — Tableau synthèse des analyses statistiques avancées (bivariées)**

Analyse	Définition	Interprétation	Exemple (issu d'un rapport d'évaluation)	Quand l'utiliser (préalables) ?
Test d'indépendance du Khi-carré ( $\chi^2$ )	Permet de tester l'indépendance entre deux variables et de déterminer l'existence (ou non) d'une relation entre deux variables.	Plus le $\chi^2$ est grand, plus la différence est grande entre la fréquence anticipée et la fréquence observée (Fox, 1999). Cela indique également une relation qui est plus forte.	<p>Dans le cadre d'une évaluation de politiques publiques visant la lutte contre les gangs de rue au Québec, Canada (Lamari Moktar <i>et al.</i>, 2015), un test de Khi-carré a été réalisé entre les variables :</p> <ul style="list-style-type: none"> <li>▪ Types d'organisations (organismes communautaires, ministères, réseau public, corps policiers, universités et autres).</li> <li>▪ Le fait d'avoir consulté le site Internet du ministère de la Sécurité publique portant sur les gangs de rue (oui, non).</li> </ul> <p>Résultats du test Khi-carré : <math>\chi^2=11,849</math>; dl=5; Sig=0, 05. Les résultats de test de Khi-carré nous permettent de dire qu'il y a une association statistiquement significative au seuil de 5 % entre les deux variables mises en relation.</p>	S'utilise seulement avec des variables catégorielles. On l'utilise pour mesurer l'association, pas la causation. Il faut deux catégories ou plus pour chaque variable.

Analyse	Définition	Interprétation	Exemple (issu d'un rapport d'évaluation)	Quand l'utiliser (préalables) ?
Coefficient de corrélation	Sert à mesurer l'intensité et le sens d'une relation entre des variables. Il s'agit d'une « mesure du degré de concentration des cas le long de la droite de régression » (Fox, 1999).	Le coefficient peut varier entre -1 et 1. Plus le résultat se rapproche d'un entier (positif ou négatif), plus la relation est forte (parfaite). Lorsque l'on obtient 0, il n'y a pas de relation (Fox, 1999).	L'étude ETIC avait pour but d'évaluer un programme d'éducation thérapeutique « en soins primaires sur la qualité de vie ». Tout d'abord, on visait à vérifier s'il y avait une corrélation à l'inclusion entre l'échelle de qualité de vie SF-36 et l'échelle Minnesota. Ensuite, il s'agissait de déterminer s'il y avait une relation entre les caractéristiques des patients (variable indépendante) et le score des échelles de qualité de vie (variable dépendante). Finalement, il existait une relation entre les deux échelles dans tous les groupes. Les scores de l'échelle SF variaient selon le genre et chez les personnes diabétiques, chez qui il était très bas. Les autres caractéristiques n'avaient pas d'impact sur le score (Guibot, 2011).	Il est utilisé pour évaluer une relation entre deux variables à intervalles ou de ratio.
Analyse de la variance ANOVA	Permet de comparer des moyennes entre plusieurs groupes. Elle permet ainsi de « tester l'existence d'un lien de causalité entre une ou plusieurs variables explicatives [...] et une variable à expliquer » (Delacroix <i>et al.</i> , 2021). L'analyse de variance permet de confronter les données d'une variable quantitative aux données d'une variable qualitative comportant plusieurs modalités.	Le postulat d'homogénéité doit être vérifié avant d'interpréter ANOVA. La valeur du F permet de vérifier s'il y a inégalité des moyennes.	Dans le cadre d'une évaluation de politiques publiques visant la lutte contre les gangs de rue au Québec, Canada (Lamari Moktar <i>et al.</i> , 2015), une analyse de la variance (ANOVA) à un facteur a été faite. Il s'agissait d'une analyse de comparaison entre un Indice d'Efficacité de Transfert de Connaissance (IETC) moyen en fonction des supports (Portail, BIP, infolettre, formations de l'axe 4, réseaux d'échange). Le résultat d'ANOVA a donné les moyennes suivantes pour chaque catégorie : Portail : 8,06 BIP : 7,92 Infolettre : 7,69 Formations : 8,82 Réseau d'échanges : 9,01 Sig : 0,0000 IETC moyen est différent en fonction des supports. Plus élevé pour les réseaux d'échange et moindre pour l'infolettre. Ce résultat est statistiquement très significatif au seuil de 1 % (p=0,00).	Lorsque l'on confronte une variable quantitative à une variable qualitative (nominale ou ordinale). La population dont est tiré l'échantillon doit être normalement distribuée. À utiliser lorsque la variable indépendante est non dichotomique. La variable dépendante doit être à intervalles ou de ratio. L'échantillon doit être aléatoire (Fox, 1999).

Analyse	Définition	Interprétation	Exemple (issu d'un rapport d'évaluation)	Quand l'utiliser (préalables) ?
Test T de Student	Sert à comparer deux modalités d'un même facteur (Delacroix <i>et al.</i> , 2021).	$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_{\bar{x}_1 - \bar{x}_2}}$ <p>Le t est la différence entre les moyennes de l'échantillon et la différence entre les moyennes de la population. La taille du t révèle donc « à combien d'erreurs types notre différence se situe par rapport à zéro » (Fox, 1999).</p>	Dans le cadre de l'analyse des effets d'une réforme organisationnelle sur le niveau de stress des employés, une comparaison des moyennes du niveau de stress avant la réforme (M=1,75) et après la réforme (M=1,87) donne une valeur du T de Student de -2,97 et sig <0,05. Nous pouvons donc conclure qu'il y a une augmentation significative du stress des employés à la suite de la réforme en moyenne de 0,12 (Ouédraogo, 2020).	À utiliser avec des variables à intervalles. L'échantillonnage doit être aléatoire.
Régression linéaire	Sert à « prévoir la valeur d'une variable en fonction de la valeur d'une autre variable » (IBM, 2021). Il s'agit du calcul d'une droite qui spécifiera la valeur que prendra une variable dépendante quand la variable indépendante variera.	Dans le graphique avec la droite créée, l'axe des abscisses présentera les valeurs de la variable indépendante, et l'ordonnée, les valeurs de la variable dépendante.	<p>Dans le cadre d'une évaluation d'impact réalisée à partir d'une approche expérimentale au Burkina Faso, l'impact des médias de masse sur l'utilisation des méthodes contraceptives modernes a été mesuré à travers le modèle de régression suivante :</p> $Y = a + bTreat + cX$ <p>Y étant la variable dépendante ou résultat.            Treat étant un indicateur d'affectation aléatoire au groupe de traitement et de contrôle.            X étant les variables indépendantes a,b,c coefficients.</p> <p>Des résultats de l'évaluation, il ressort que l'utilisation de la contraception moderne a augmenté de 5,6 points de pourcentage dans le groupe de bénéficiaires comparativement au groupe de contrôle. Ce résultat est statistiquement significatif au seuil de 5 % (p=0,046) (Rachel <i>et al.</i>, 2021).</p>	Les valeurs doivent être normalement distribuées. Les variances doivent être homogènes.

Il est à noter que l'évaluateur ou l'évaluatrice n'a pas nécessairement besoin de maîtriser la technique du procédé lui-même, mais plutôt d'être en mesure de commander différentes analyses à un professionnel en statistique et de pouvoir interpréter les résultats de ces analyses pour répondre aux questions d'évaluation.

## **Intégration des données qualitatives et quantitatives en fonction des finalités d'une stratégie mixte de collecte et d'analyse de données**

L'intégration d'une phase qualitative à une phase quantitative (ou l'inverse) dans la phase d'analyse s'inscrit à plusieurs objectifs. Généralement, on compte trois types d'intégration en méthodes mixtes, qui ont chacune leurs visées (complémentarité, tension dialectique et unification) (Pluye, 2019). Un premier type d'intégration est la connexion de phases ou l'analyse mixte séquentielle. Dans le cadre de ce type d'intégration, les résultats de la première phase (quantitative ou qualitative) informent la deuxième (Pluye, 2019). Ensuite, on retrouve la comparaison de résultats, qui décrit un processus où la collecte et l'analyse des deux phases se font séparément, pour finalement être comparées. Finalement, on retrouve l'assimilation des données, qui consiste à transformer ou à fusionner les données lors de l'analyse. Ainsi, les données qualitatives peuvent devenir quantitatives (Pluye, 2019). L'inverse est aussi possible et on peut également y aller au cas par cas. Afin que l'intégration soit réussie, il ne s'agit pas de présenter une série d'informations, mais de créer une réelle dialectique entre les phases, toujours dans le but de répondre à la question d'évaluation (Castro, 2020).

### **Considérations pratiques pour l'intégration en méthodes mixtes**

Survol des erreurs les plus fréquentes dans la production des réponses (pièges à éviter à cette étape de l'évaluation) :

- Argumentation selon un raisonnement descriptif, c'est-à-dire sans éléments de comparaison. Il est important de créer un dialogue entre les phases afin que celles-ci s'informent mutuellement.
- Absence de lien avec la question de départ. Lors de l'intégration, il faut toujours garder en tête le but de la démarche : répondre à la question d'évaluation.

Les principales qualités d'une bonne réponse :

- Utilité – en lien avec l'utilisation prévue des résultats;
- Validité – relatifs à la qualité des données et l'interprétation qu'on en fait;
- Acceptation – importance d'impliquer les parties prenantes dans l'analyse et l'interprétation des données – condition essentielle à la réception des constats et des argumentaires pour répondre aux questions d'évaluation.

Il est important de ne pas réaliser cette phase seul. L'intégration nécessite une expertise qualitative et quantitative qui exige un travail collectif.

Les possibilités d'intégration des analyses qualitatives et quantitatives demeurent intéressantes pour l'argumentaire évaluatif. Toutefois, il n'y a pas encore de consensus sur la meilleure façon de procéder. Le temps restreint de la collecte et de l'analyse des données en évaluation limite les intégrations les plus poussées.

## Conclusion

Bien que l'évaluateur ou l'évaluatrice suive le plan établi lors de la phase méthodologique, il ne devrait pas s'y limiter. Il doit plutôt rester ouvert et permettre aux constats d'émerger lorsque nécessaire selon un processus itératif entre la collecte, le traitement, l'analyse et l'interprétation. C'est souvent ce mouvement de va-et-vient qui va permettre de répondre aux questions d'évaluation. Le grand défi de la phase analytique dans la démarche évaluative demeure celui de l'intégration des données dans une perspective mixte.

## Messages clés

1. L'analyse des données qualitatives nécessite autant de rigueur que celle de l'analyse quantitative pour s'assurer de la validité des résultats. Plusieurs possibilités s'offrent au chercheur, mais les choix doivent être faits en fonction des questions d'évaluation.
2. Dans les analyses qualitatives, la codification des données peut faire émerger ou disparaître certains thèmes : l'évaluateur ou l'évaluatrice doit alors s'assurer de l'exhaustivité, de la qualité et de la pertinence de ceux-ci.
3. Dans l'analyse quantitative, il importe donc de se référer aux questions d'évaluation, aux types de variables à l'étude ainsi qu'à la taille de l'échantillon ou de la population étudiée pour déterminer le genre d'analyse qui sera à effectuer.
4. Les stratégies mixtes d'analyse des données renvoient à la notion d'intégration des données qui est une opération délicate. Les évaluatrices et évaluateurs doivent être particulièrement conscients des limites méthodologiques pour l'interprétation des données pour la réponse aux questions d'évaluation dans un contexte où le temps de collecte et d'analyse est particulièrement restreint.

## Pour aller plus loin

**Pour en savoir plus** sur les fonctions de la carte heuristique et les principes à respecter pour son élaboration.

<https://www.lescahiersdelinnovation.com/la-carte-heuristique-ou-mind-map/>

## Références bibliographiques

- Andreani, J. C. et Conchon, F. (2005). Méthodes d'analyse et d'interprétation des études qualitatives : État de l'art en marketing. *Revue française du marketing*, 201, 5-21.
- Antin, T. M. J., Constantine, N. A. et Hunt, G. (2014). Conflicting Discourses in Qualitative Research : The Search for Divergent Data within Cases. *Field Methods*, 27(3), 211-222. <https://doi.org/10.1177%2F1525822X14549926>
- Baribeau, C. et Germain, M. (2010). L'entretien de groupe : considérations théoriques et méthodologiques. *Recherches qualitatives*, 29(1), 28-49.
- Castro, D. (2020). La recherche à base de méthodologie mixte : une nécessité scientifique et professionnelle. Dans A. Schweizer, M. del Rio Carral et M. Santiago Delefosse (dir.), *Les méthodes mixtes en psychologie : Analyses qualitatives et quantitatives : de la théorie à la pratique* (p. 63-84). Dunod.
- Cherkaoui, A. et Haouata, S. (2017). Éléments de réflexion sur les positionnements épistémologiques et méthodologiques en sciences de gestion. *Revue Interdisciplinaire*, 1(2), 1-20.
- Delacroix, E., Jolibert, A., Monnot, É. et Jourdan, P. (2021). Les analyses de variance univariée et multivariée. Dans E. Delacroix, A. Jolibert, É. Monnot et P. Jourdan (dir.), *Marketing Research : Méthodes de recherche et d'études en marketing*. (p. 269-297). Dunod.
- Del Ré, A. et Hilario, R. N. (2014). Concilier les données qualitatives et quantitatives : un défi pour les recherches en acquisition du langage. *Corela – cognition, représentation, langage*, HS-15. <https://doi.org/10.4000/corela.3564>
- Delve. (2021). *Inter-coder reliability*. <https://delvetool.com/blog/intercoder>
- Fortin, M.-F. et Gagnon, J. (2016). *Fondements et étapes du processus de recherche : Méthodes quantitatives et qualitatives* (3<sup>e</sup> édition). Chenelière Éducation.
- Fox, W. (1999). *Statistiques sociales*. Les Presses de l'Université Laval.
- Glennerster, R., Murray, J. et Pouliquen, V. (2021). *The Media or the Message? Experimental Evidence on Mass Media and Modern Contraception Uptake in Burkina Faso* (Document de travail de la SCDA WPS/2021-04). Centre for the study of African Economies.
- Guibot, S. (2011). *E.T.I.C. : corrélation des résultats de deux échelles de qualité de vie et influence des caractéristiques des patients sur l'évaluation de la qualité de vie* [mémoire, Université de Clermont]. <http://www.sudoc.abes.fr/cbs/xslt/DB=2.1//SRCH?IKT=12&TRM=158155076>
- Houle, S. et Hurteau, M. (2012). La production d'un jugement crédible. Dans M. Hurteau, F. Guillemette et S. Houle (dir.), *L'évaluation de programme axée sur le jugement crédible* (p. 175-192). Presses de l'Université du Québec.
- Houle, S., Hurteau, M. et Marchand, M.-P. (2012). Les processus de production et de crédibilisation du jugement en évaluation. Dans M. Hurteau, F. Guillemette et S. Houle (dir.), *L'évaluation de programme axée sur le jugement crédible* (p. 77-99). Presses de l'Université du Québec.
- IBM. (2021). *Régression linéaire*. <https://www.ibm.com/fr-fr/analytics/learn/linear-regression>

- Lamari, M., Antett, E., Sauvain, V., Jacob, J. et Viel, A. (2014). *Recherche évaluative sur la mise en œuvre, les effets et les retombées du Plan d'intervention québécois sur les gangs de rue 2011 – 2014. Rapport final*. Centre de recherche et d'expertise en évaluation. École nationale d'administration publique.
- Mucchielli, A. et Paillé, P. (2021). *L'analyse qualitative en sciences humaines et sociales* (5<sup>e</sup> édition). Armand Colin.
- Nations Unies. (2004). *Études d'évaluation thématique : approche qualitative de la collecte de données (module 6)*. [https://www.unodc.org/documents/GAP/GAP%20toolkit%20module%206%2003-90135\\_FRENCH.pdf](https://www.unodc.org/documents/GAP/GAP%20toolkit%20module%206%2003-90135_FRENCH.pdf)
- Nowell, L. S., Norris, J. M., White, D. E. et Moules, N. J. (2017). Thematic Analysis : Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, 16, 1-13. <https://doi.org/10.1177/1609406917733847>
- Ouédraogo, M. (2020). *Cours de méthodes statistiques pour évaluation de programmes*. Automne 2021. École nationale d'administration publique.
- Patton, M. (2008). *Utilization-Focused Evaluation* (4<sup>e</sup> édition). Southbound.
- Pluye, P. (2019). L'intégration en méthodes mixtes. Dans V. Ridde et C. Dagenais (dir.), *Évaluation des interventions de santé mondiale*. IRD Éditions.
- Taylor-Powell, E. et Renner, M. (2003). *Analyzing qualitative data. Program Development & Evaluation*. University of Wisconsin. <https://cdn.shopify.com/s/files/1/0145/8808/4272/files/G3658-12.pdf>
- Turcotte, D., Dufour, I. F. et Saint-Jacques, M.-C. (2009). Les apports de la recherche qualitative en évaluation de programmes. Dans M. Alain et D. Dessureault (dir.), *Élaborer et évaluer les programmes d'intervention psychosociale* (p. 195-216). Presses de l'Université du Québec.





# Partie IV

## Phase de diffusion et de transfert



# La présentation des résultats d'évaluation

*Abdoul Aziz Sall*  
*Élodie C. Roy*



## Introduction

La présentation des résultats de l'évaluation intervient après la collecte et l'analyse des données. Elle est d'une importance majeure, car l'utilisation des résultats par les parties prenantes est en partie tributaire d'une présentation et d'une communication efficaces. Plusieurs théoriciens, auteurs et évaluateurs (Alkin, 1991; King, 1998; Stufflebeam, 2001; Patton, 2008; Preskill et Torres, 2001; Wholey, 1981, 2001) considèrent d'ailleurs que c'est l'utilisation des résultats qui doit orienter le processus et donner du sens à l'évaluation. Selon Patton, très connu pour son rôle dans le développement de l'évaluation axée sur l'utilisation (*utility focused evaluation*), « les évaluations devraient être jugées par leur utilité et leur véritable usage » (Patton 2008, p. 37). Présenter les résultats, c'est donc en quelque sorte créer les conditions propices à une prise de décision éclairée.

Présenter les résultats peut se diviser en deux grands axes : un en lien avec le fond (c'est-à-dire le contenu) et l'autre avec la forme (le mode de présentation). Dans les deux cas, l'appropriation et l'utilisation des résultats par les parties prenantes comptent parmi les finalités. Au niveau du fond, la présentation des résultats (soit les connaissances produites et les données générées par l'évaluation) devrait aider les parties prenantes à comprendre comment les constats de l'évaluation ont été obtenus (par exemple, en expliquant les détails du processus d'évaluation, les méthodes de collecte et d'analyse utilisées) en plus d'aider à révéler les tendances qui ont été mises au jour grâce à la collecte et à l'analyse des données réalisées. Dans le contexte de la « révolution des données » (IEAG, 2014), la visualisation efficace des données joue de plus en plus un rôle important à cet égard, notamment pour aider à les rendre plus digestes, mais aussi pour mettre en évidence les données et les statistiques. Au niveau de la forme, la présentation et la diffusion des résultats peuvent s'articuler autour de divers moyens (par exemple, le rapport écrit, la présentation orale, etc.) et impliquer les parties prenantes à différents stades afin d'assurer une communication efficace. Bien que de nouvelles façons de présenter les résultats fassent leur apparition, ultimement, le rapport écrit demeure, à ce jour, le moyen de présentation le plus commun.

## Les moyens de présentation des résultats

Présenter les résultats revient généralement à exposer sous forme de rapport écrit les connaissances produites dans le cadre de cet exercice d'appréciation et de jugement qu'est l'évaluation. Le rapport écrit n'est toutefois pas la seule façon de présenter les résultats. Il existe différentes façons de communiquer les résultats de l'évaluation, qui vont des plus traditionnelles – parmi lesquelles on peut citer les rapports écrits et les présentations orales – aux plus innovantes – par exemple : les comptes rendus créatifs et les comptes rendus utilisant des formats électroniques. Il est possible dans certains cas de combiner plusieurs moyens de présentation pour une seule évaluation. Le choix de la forme

de présentation des résultats est souvent dicté par le client qui demande l'évaluation, le moment de la présentation et la stratégie de communication de l'évaluation (ce dernier aspect sera examiné sous l'angle du transfert de connaissances dans le chapitre 19 : *Un processus de transfert de connaissances pour favoriser l'utilisation des résultats d'évaluation*).

L'utilisation la plus diversifiée et la plus inventive des formes de présentation des résultats se trouve souvent chez les organisations internationales et les organisations non gouvernementales (ONG). En plus des rapports écrits, enrichis de moyens novateurs de visualisation et de mise en relief des données et informations importantes, certaines de ces organisations ont recours à d'autres façons de présenter les résultats comme des présentations orales, des comptes rendus créatifs et des comptes rendus électroniques. Plusieurs raisons peuvent expliquer ce fait. S'agissant de l'évaluation, ces organisations ont un penchant pour les approches participatives, ce qui élargit et diversifie le champ des acteurs concernés par l'évaluation et pour lesquels il y a un intérêt à présenter les résultats de manière différenciée et adaptée. Ces organisations sont plus à la pointe des nouvelles pratiques que sont la gestion axée sur les résultats (GAR) et l'évaluation, comme en témoignent leurs nombreuses productions (guides, manuels, notes techniques, lignes directrices, bonnes pratiques, etc.) qui font évoluer qualitativement ces pratiques. L'utilisation des moyens électroniques de communication, notamment les médias sociaux, prend de plus en plus d'ampleur dans les pratiques de ces organisations. Il est courant de voir afficher les logotypes des grands médias sociaux (Facebook, Twitter, Instagram, LinkedIn, etc.) sur les documents officiels de ces organisations pour montrer leur présence dans ce nouvel espace médiatique et pour renvoyer aux contenus qu'elles y produisent et présentent.

Au sein de l'administration publique, quand l'évaluation est menée, la présentation des résultats se fait quasi exclusivement par le moyen du rapport écrit d'évaluation au détriment des autres moyens plus innovants de présentation des résultats. Plusieurs raisons peuvent expliquer ce fait. Il survit, dans une certaine mesure, dans les administrations publiques, l'évaluation de type managérial (Baron et Monnier, 2003), ce qui restreint considérablement le champ de diffusion des résultats. La culture de l'écrit, avec les canaux consacrés de diffusion de l'information administrative, reste une des traditions de l'Administration les mieux conservées. En matière de gestion par résultats dans les administrations publiques, il y a un fossé entre les « espoirs » et la « réalité », entre « le discours » et les « réalisations », et la culture du résultat n'y est souvent pas encore suffisamment ancrée, au point d'apporter des innovations majeures dans la présentation et la mise en relief des résultats (Emery, 2005). Enfin, l'utilisation des nouveaux moyens de communication et de diffusion de l'information reste timide dans les communications gouvernementales. De manière générale, en matière de transformation numérique, l'administration publique a des difficultés à s'adapter aux nouvelles avancées (Caron, 2019).

### **Différentes façons de présenter les résultats**

Le tableau ci-dessous présente de façon succincte certains avantages et inconvénients liés à différents moyens de présentation et de diffusion des résultats de l'évaluation, ainsi que des liens hypertextes vers des exemples concrets qui peuvent servir d'inspiration.

Peu importe les avantages, les limites ou les préférences de l'évaluateur ou évaluatrice ou de l'équipe d'évaluation envers certaines approches plutôt que d'autres, le format, les moyens de diffusion et les stratégies de communication seront toujours à déterminer avec le client principal avant de commencer l'évaluation.

**Tableau 1 — Différentes façons de communiquer les résultats d'évaluation.**  
 Inspiré de Stetson, 2008 et Duarte, 2016.

Moyen	Avantages	Inconvénients / Limites	Exemples
Rapport écrit	Donne accès aux mêmes informations à tous. Les utilisateurs peuvent prendre connaissance de l'information à leur propre rythme et y revenir au besoin.	Les contenus denses et très textuels rebutent parfois l'audience. Le temps de lecture pour les parties prenantes s'avérera plus ou moins long dépendamment de la longueur du rapport. Le style d'écriture, la précision et la clarté du langage utilisé peuvent avoir une influence sur la compréhension. Contrairement à la communication orale, le rapport écrit ne permet pas aussi facilement les échanges. Le temps pour la préparation et la rédaction est souvent important.	<u>Évaluation du projet de renforcement du contrôle citoyen et de la redevabilité sociale dans 49 communes au Burkina Faso</u>
Présentation orale (par exemple : panel, radio, balado ( <i>podcast</i> ), télévision, communication informelle, téléconférences, vidéoconférences, médias)	La communication verbale permet de présenter un message plus spécifique, personnalisé et interactif par rapport au rapport écrit. Cette façon offre aussi la possibilité d'interagir avec l'auditoire.	Les présentations orales supportent difficilement le partage de l'information sans la présence du présentateur ou d'un narratif qui accompagne chaque diapositive. La communication est influencée par le volume, la vitesse et la clarté de la parole. Elles se font également à un moment précis, ce qui peut limiter la participation en fonction des disponibilités et des conflits d'horaire des parties prenantes. Le temps de préparation et d'organisation peut être important.	<u>Présentation du Rapport d'évaluation externe du projet IFADEM au Tchad</u>
<i>Slidedocs</i>	À la croisée entre rapport écrit et présentation orale, les <i>Slidedocs</i> sont des documents développés avec un logiciel de présentation, mais qui sont faits pour être lus au lieu d'être projetés sur un écran pour en faire un contenu plus facilement assimilable et partageable.	La création des <i>Slidedocs</i> nécessite un certain savoir-faire au niveau de la conception ( <i>design</i> ) et le temps de préparation et de réalisation peut être non négligeable.	<u>Évaluation de l'égalité des genres et du renforcement du pouvoir des femmes et des filles au Moyen-Orient et au Maghreb 2015-16 à 2019-2020</u>
Comptes rendus créatifs (par exemple : vidéo ou pièces de théâtre)	En présentant une synthèse créative de l'évaluation, les comptes rendus créatifs peuvent contribuer à maintenir l'intérêt. Ces formats peuvent également être utiles quand il s'agit de transmettre des informations à des parties prenantes analphabètes ou pour qui la tradition orale est importante.	Ces moyens peuvent engendrer un phénomène de « spectateur passif » et offrent souvent moins de possibilités d'interagir avec les participants que lors des présentations orales. En s'apparentant à un divertissement, l'utilisateur peut ne pas réfléchir en profondeur au contenu et aux messages qui sont livrés. Ces moyens nécessitent également un savoir-faire particulier en complément des capacités pour l'évaluation et peuvent entraîner des coûts importants.	<u>Évaluation <i>ex post</i> écrite et filmée du projet de développement du parc national du Limpopo</u>

Moyen	Avantages	Inconvénients / Limites	Exemples
Communications électroniques (par exemple : brèves communications ou bulletins présentés sur les sites Internet ou les médias sociaux, affiches)	Les communications électroniques permettent de présenter un message court, facile à lire et avec un visuel attrayant. Ces formats sont souvent avantageux du point de vue visuel grâce à l'utilisation de la couleur, de la mise en page, de titres variés et des illustrations.	Ce genre de communication limite l'accès à l'information à ceux qui ont un ordinateur avec Internet. Le caractère informel de certaines plateformes peut laisser l'impression que l'information partagée est non officielle ou moins crédible. De plus, ces formats peuvent parfois être moins faciles à utiliser quand l'évaluation est surtout qualitative et quand une description complète du contexte de l'évaluation est cruciale pour interpréter les résultats.	<p><a href="#">Rapport d'évaluation de la politique scientifique de la Wallonie-Bruxelles 2018-2019</a></p> <p><a href="#">Atlas 2020 des Objectifs de développement durable: des visualisations de données originales pour mieux comprendre les enjeux (worldbank.org)</a></p> <p><a href="#">Affiche sur l'enseignement et l'évaluation des attitudes dans un programme collégial</a></p>

## Le rapport d'évaluation

Il n'existe pas une seule et unique façon de rédiger un rapport d'évaluation. Toutefois, il existe certaines façons de faire, des normes et des standards qui sont de plus en plus admis dans la profession. La recherche de l'utilisation réelle de l'évaluation par les parties prenantes passe entre autres par la capacité, à travers le rapport, d'identifier, de mettre en relief et de résumer les informations, les données, les connaissances et les résultats essentiels, et de convaincre de la rigueur de la démarche et des méthodes choisies pour l'évaluation. Tout cela requiert des aptitudes d'analyse et de synthèse, de la précision et de la concision. Les composantes suivantes peuvent servir de point de départ pour structurer le rapport d'évaluation.

- **Un résumé exécutif ou sommaire exécutif**

Puisqu'il est admis que les résultats de l'évaluation doivent fournir aux décideurs des informations et des données factuelles/probantes concernant les performances et les bonnes pratiques pour permettre une prise de décision éclairée et alimenter les décisions, il est nécessaire de mettre en œuvre une démarche efficace de présentation des résultats à même de faciliter la prise de décision. La rédaction du résumé exécutif participe à cette stratégie qui tient compte du temps limité dont disposent les décideurs pour exploiter l'ensemble du contenu d'un rapport écrit et qui permet de rassembler tous les éléments nécessaires à la prise de décision. De ce point de vue, le résumé exécutif doit contenir tous les éléments nécessaires à la prise de décision et être autoporteur dans son contenu pour le décideur, tout en demeurant court (généralement entre 2 et 5 pages).

De plus, il peut contenir, pour rappel, les éléments essentiels de la phase de conception de l'évaluation, par exemple les questions d'évaluation, la méthodologie utilisée, en plus de présenter un résumé des constats, des conclusions et des recommandations (s'il y a lieu). En ce sens, le résumé exécutif diffère



de l'introduction du rapport d'évaluation. Il peut inclure, en plus du texte, des éléments de visualisation qui permettent de mettre en relief certaines données et informations d'importance particulière et ainsi faciliter un survol efficace de l'ensemble du contenu du rapport.

- **L'introduction**

L'introduction doit être rédigée de façon à capter l'attention du lecteur. Pour ce faire, l'évaluateur ou l'évaluatrice doit user de techniques d'écriture efficaces à cet effet dès les premières phrases. Le démarrage par un « crochet » (phrase-choc, *punch line*) est l'une de ces techniques. Elle doit essentiellement contenir le but de l'évaluation, qui doit être clairement défini, notamment la (ou les) raison pour laquelle cette évaluation s'impose à ce moment précis, ainsi que les informations recherchées et les éléments clés du contexte de l'évaluation (Sall, 2019). Il n'est toutefois pas utile de donner trop de détails sur l'intervention ou l'objet d'évaluation à ce stade.

- **La méthodologie de l'évaluation**

La méthodologie de l'évaluation doit être présentée de façon à démontrer au lecteur la rigueur scientifique avec laquelle l'évaluation a été menée. Les éléments qui attestent de la rigueur scientifique d'une évaluation sont essentiellement les suivants : le cadre théorique (modèle logique, la chaîne des résultats ou encore la théorie du changement de l'intervention), les questions d'évaluation, la matrice d'évaluation et les méthodes de collecte et d'analyse utilisées. Il est important de signaler les limites de la démarche méthodologique. Il n'y a pas de méthodologie sans limites objectives. Le fait de mentionner ces limites renforce la rigueur scientifique de la démarche d'évaluation.

### **Encadré 1 – À propos des limites et difficultés de l'évaluation**

S'il n'existe pas de démarche méthodologique sans limites objectives, c'est parce qu'il n'existe pas de choix technique parfait pour mener une évaluation, et il n'existe pas non plus d'évaluation qui ne présente pas de difficultés dans ses aspects pratiques et opérationnels. Les limites qui relèvent des choix techniques pour mener une évaluation ne doivent pas être perçues comme des lacunes ou des insuffisances, mais plutôt comme des bornes, des délimitations nécessaires entre différents choix possibles. De ce point de vue, le fait de signaler une limite inhérente à un choix revient tout simplement à dire jusqu'où ce choix peut conduire dans la démonstration des résultats d'une évaluation et implicitement à dire tout ce que ce choix ne permet pas de faire. Il est nécessaire de souligner la différence entre ce type de limites et les difficultés que l'on peut rencontrer sur le terrain lors d'une évaluation et qu'il faut aussi nécessairement signaler. Cela peut aller des événements inattendus qui ont pu remettre en question le temps ou les ressources qui étaient prévues, des problèmes d'accès aux sources d'information, à la qualité de certaines données découvertes. Pour l'évaluateur ou l'évaluatrice, le fait de signaler les limites et les difficultés rencontrées lors d'une évaluation, c'est satisfaire à un principe éthique et de transparence qui le lie en tant que professionnel. Le lecteur qui a connaissance des limites et des difficultés d'une évaluation détient une bonne partie des éléments qui peuvent éclairer son appréciation de la qualité du travail d'évaluation réalisé.

- **Les constats et les résultats**

Les constats et les résultats englobent ici les connaissances produites par l'évaluation ainsi que les effets (c'est-à-dire les modifications ou les changements démontrés ou observés) causés par l'intervention évaluée. Ils permettent de donner du sens à l'évaluation, ce qui revient à répondre aux questions d'évaluation en définitive. La partie consacrée à l'organisation des résultats est une des composantes majeures du rapport d'évaluation. Sa présentation requiert des précautions particulières. Les constats et les résultats doivent être accessibles au lecteur, donc compréhensibles (à ce niveau, les analyses techniques des données ne sont pas toujours pertinentes). Ils doivent uniquement porter sur les résultats importants. Ils doivent être organisés de façon à répondre séparément aux questions d'évaluation (qui découlent des critères d'évaluation). Ils doivent être présentés de façon à attirer l'attention du lecteur, notamment par l'utilisation d'encadrés dans le texte, de tableaux et de graphiques. Des exemples de visualisations qui peuvent être utilisées dans les rapports d'évaluation sont d'ailleurs présentés dans la sous-section suivante intitulée « La visualisation des résultats ».

### **Encadré 2 – Attirer l'attention du lecteur sur les éléments importants dans un rapport**

Dans cet extrait d'un rapport d'évaluation, des titres et des intertitres rappellent les critères d'évaluation pour faciliter le repérage de l'information. De plus, pour attirer l'attention des lecteurs, l'équipe d'évaluation a utilisé des caractères gras afin de faire ressortir les conclusions de son analyse. Les encadrés et les puces sont d'autres moyens qui peuvent être utilisés pour attirer le regard et orienter la lecture vers ce qui devrait intéresser les parties prenantes et pour mettre en relief les points importants au sein du rapport.

#### **Cohérence**

Le critère de cohérence a été analysé en fonction de la cohérence interne – qui concerne les synergies et les liens entre les interventions et programmes de l'organisation – et la cohérence externe, qui a examiné spécifiquement le rôle joué par l'organisation dans l'environnement international. [...]

#### **Cohérence externe**

**L'organisation n'a toujours pas de démarche systématique pour favoriser le développement et le maintien de partenariats stratégiques.** Ces préoccupations ne sont pas nouvelles; le rapport d'évaluation de 2018 notait que les partenariats étaient à la fois « la principale force de l'organisation, mais aussi un domaine de vulnérabilité ». Ceci reste largement vrai à ce jour.

Traduction libre de : C. Voyadzis et E. C. Roy, 2021

- **La conclusion**

Les éléments de la conclusion doivent être en lien avec les questions d'évaluation, les objectifs ou la finalité de l'évaluation. La conclusion doit nécessairement s'appuyer sur les constats de la partie « résultats ». Elle doit revenir de manière succincte sur les éléments clés de l'évaluation. La conclusion n'est pas un résumé du rapport d'évaluation. Elle n'est pas non plus le lieu de discussions théoriques nouvelles ou d'ouverture sur d'autres questions d'évaluation, comme cela se fait souvent dans le cadre de recherches scientifiques.

- **Les recommandations**

Au sein de la communauté des évaluateurs et évaluatrices, il existe principalement deux écoles de pensée en ce qui concerne les recommandations : l'école qui prône la formulation de recommandations au sein du rapport d'évaluation, et l'école qui prône la non-formulation de recommandations dans un rapport d'évaluation. Dans la pratique, en particulier dans les administrations publiques, la formulation de recommandations est de rigueur. Quand elles sont formulées, elles doivent être en cohérence avec les résultats et les constats soulevés dans la conclusion. Elles doivent également être formulées de façon à faciliter leur application (recommandations clés, quoi faire, quand et comment, présenter des scénarios différents, etc.). Le chapitre 18 : *L'élaboration des recommandations pour favoriser l'action* aborde de manière plus approfondie les enjeux de la présentation de recommandations aux parties prenantes.

- **Les annexes**

Les annexes concernent l'ensemble des documents et outils qui permettent de comprendre comment l'évaluation a été menée. Celles-ci peuvent inclure par exemple : les termes de référence de l'évaluation, les outils de collecte de données, la liste des personnes interrogées et des sites visités, des informations complémentaires sur la méthodologie utilisée (par exemple, le calcul des échantillons).

### Encadré 3 – Consignes générales pour un rapport écrit de qualité

- Aller à l'essentiel et s'adapter au public visé;
- Utiliser un vocabulaire simple, actif, positif et sensible sur le plan culturel;
- Éviter les abréviations et les acronymes;
- Limiter les détails sur le contexte;
- Fournir suffisamment d'informations sur la méthodologie de façon que les lecteurs puissent juger de sa crédibilité;
- Organiser le rapport autour des principaux thèmes ou autour des questions d'évaluation;
- Présenter les points majeurs en premier ou les mettre en évidence;
- Appuyer les conclusions et les recommandations avec des preuves/des données;
- Joindre les informations techniques (instruments de collectes de données, résultats plus longs) en annexe;
- Prévoir assez de temps pour valider et réviser l'information et, si possible, faire relire le rapport par une personne indépendante.

## La visualisation des résultats

La présentation des résultats vise à mettre à la disposition des parties prenantes des données probantes afin que celles-ci contribuent à la réflexion et à la prise de décision. Pour faciliter la communication des informations issues des évaluations, il existe une large gamme de visualisations qui peuvent être utilisées, selon le type d'utilisateur ou encore du type de données à visualiser (qualitatives ou quantitatives).

La visualisation des résultats est souvent utilisée pour :

- Augmenter l'intérêt ou attirer l'attention du lecteur sur certains éléments en particulier;
- Communiquer l'information de façon plus claire et plus facilement compréhensible que par du texte;
- Améliorer l'impact d'un message ou du rapport en entier.

Dans le cadre d'une évaluation, la démarche de présentation des résultats peut suivre une logique du type : visualisation, interprétation, appréciation.

1. La visualisation de données est la représentation visuelle de l'information. Cette façon de présenter les résultats (ou les constats principaux) aide à rendre les importantes quantités d'informations, souvent issues de l'évaluation, plus digestes pour les utilisateurs. En effet, le fait de disposer de données brutes compilées au sein de longs tableaux ne met pas toujours les utilisateurs et les parties prenantes dans une position idéale ou même favorable pour faire de bons choix et prendre de bonnes décisions. Une bonne visualisation des données aide à transformer des chiffres ou des données brutes en éclairage significatif sur des questions d'importance. Les informations et les résultats présentés visuellement pourront parfois être mieux compris, ce qui facilitera leur utilisation.
2. Pour favoriser la compréhension, il est important de ne pas seulement présenter l'élément visuel, mais également de l'interpréter. Par exemple, il pourra être utile de spécifier le contexte spécifique qui a donné lieu aux résultats, les raisons qui ont motivé certains choix, etc. En somme, il s'agit ici d'offrir aux utilisateurs l'information contextuelle qui leur permettra de bien cerner les enjeux qui sous-tendent les résultats présentés.
3. Finalement, la visualisation et l'interprétation doivent être accompagnées d'une appréciation (et des motivations/arguments qui soutiennent cette appréciation), car le jugement sur la valeur ou le mérite est à la base de ce qu'est la profession d'évaluation.

**Encadré 4 – Exemple : visualisation, interprétation, appréciation***Visualisation*

Normes Cliniques	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

**Légende :** Satisfait – vert | Non satisfait – blanc (vide)

**Figure 1**

Niveau de respect des normes de soins dans l'échantillon (10) de cliniques rurales du district du Nord-Est du pays visité, 2010.

*Interprétation*

Afin de déterminer dans quelle mesure les services prénataux fournis par les cliniques rurales du district du Nord-Est du pays sont adéquats, l'équipe d'évaluation a visité un échantillon aléatoire de 10 cliniques rurales dans cette zone. Lors de ces visites, l'équipe d'évaluation a pu s'entretenir avec des professionnels de la santé travaillant dans les cliniques, et observer leurs pratiques. De plus, une enquête menée auprès de femmes qui ont utilisé les services prénataux de la clinique au cours l'année précédente a permis d'obtenir des informations additionnelles sur les services offerts dans les cliniques visitées.

*Appréciation*

À la suite de l'analyse des données collectées, l'équipe d'évaluation estime que de manière générale, les services prénataux fournis par les cliniques rurales du district du Nord-Est du pays sont inadéquats. En effet, comme démontré dans la figure 1 ci-dessus, seulement 2 des 10 cliniques rurales visitées satisfont aux six normes de soins établies, tandis que les 8 autres (80 %) n'ont pas atteint au moins trois de ces normes. La norme la moins fréquemment respectée (7 cliniques, soit 70 % des cas) était le « maintien d'un ratio minimum entre le personnel et les patients » (norme 1).



Traduction libre et adaptation de : USAID, 2010




Suivant le type de résultats à présenter, certaines visualisations seront plus appropriées et mieux adaptées que d'autres. L'information qui suit est loin d'être exhaustive, mais elle vise quand même à donner un aperçu de la vaste gamme de façons possibles de présenter les résultats de manière visuelle en plus d'offrir quelques astuces pour bien utiliser chaque moyen.

## Moyens quantitatifs de visualisation des résultats

Il existe plusieurs types différents de diagrammes qui permettent de visualiser les données quantitatives. Quelques diagrammes et graphiques utilisés fréquemment dans les rapports sont présentés dans le tableau ci-dessous avec des liens hypertextes vers des exemples issus de rapports réels.

**Tableau 2 — Exemples de diagrammes et graphiques fréquemment utilisés pour visualiser des données quantitatives. Contenu inspiré de : Statistique Canada, *Les statistiques : le pouvoir des données! Types de diagrammes : Utilisation des diagrammes* (statcan.gc.ca), Stetson, 2008, Ferdio Data Viz Project. Icônes par : Anconer Design et Sam Smith, The Noun Project.**

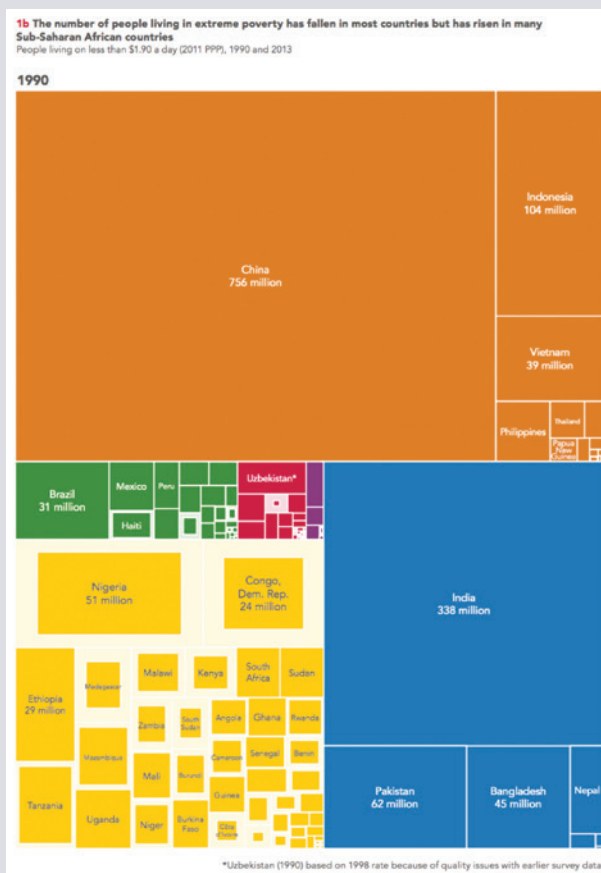
Icône	Brève description	Usages appropriés (✓) et à éviter (X)	Lien vers des exemples
	<p>Les <b>diagrammes à bandes</b> sont composés d'un axe et d'une série de bandes horizontales ou verticales. Plus les bandes sont longues ou hautes, plus leur valeur est grande.</p>	<p>✓ : Ces types de diagrammes fonctionnent bien pour comparer des valeurs, notamment pour les séries chronologiques ou différentes catégories/éléments.</p> <p>Ils sont un excellent choix pour mettre l'accent sur un changement d'amplitude.</p> <p>X : Si vous voulez comparer plus de neuf ou dix éléments, il pourrait être préférable d'utiliser un diagramme linéaire. Les diagrammes à bandes superposés peuvent aussi parfois être difficiles à lire.</p>	<p><a href="#">OMD-Agenda 2063/ODD–Rapport de transition 2016</a></p> <p><b>Vertical :</b> Figure 1.3., p. 7 Figure 1.5, p. 10</p> <p><b>Horizontal :</b> Figure 1.7, p. 12</p> <p><b>Bandes superposées :</b> Figure 2.1, p. 20 Figure 2.2., p. 21</p>
	<p>Les <b>diagrammes linéaires</b> représentent des séries de points de données appelés « marqueurs » reliés par des segments de ligne continue.</p>	<p>✓ : Ces types de diagrammes sont utiles pour révéler des tendances dans le temps, des mouvements, des cycles ou des relations entre les données, et pour comparer les tendances de différents groupes de variables.</p> <p>X : Évitez de montrer plus de trois ou quatre lignes par diagramme, car cela peut rendre la lecture et la compréhension du graphique plus difficile.</p>	<p><a href="#">OMD-Agenda 2063/ODD–Rapport de transition 2016</a></p> <p><b>Linéaire multiple :</b> Figure 4.5, p. 49 Figure 8.1., p. 91</p>

Icône	Brève description	Usages appropriés (✓) et à éviter (X)	Lien vers des exemples
	<p>Les <b>diagrammes d'aires</b> sont très similaires aux diagrammes linéaires, mais l'aire comprise sous la ligne est remplie, indiquant un volume.</p>	<p>✓ : Ces types de diagrammes sont utiles pour visualiser les différences d'amplitudes entre des phénomènes.</p> <p>X : Il est important de bien régler le contraste entre les aires pour augmenter la lisibilité.</p>	<p><a href="#">OMD-Agenda 2063/ ODD – Rapport de transition 2016</a> Figure 8.2., p. 92</p>
	<p>Les <b>diagrammes circulaires</b> (parfois aussi appelés : pointes de tarte, camemberts ou même beignes) sont formés d'un cercle divisé en segments.</p>	<p>✓ : Les diagrammes circulaires montrent habituellement les parties d'un tout. Un segment peut être séparé du reste du cercle afin de souligner l'importance de l'information.</p> <p>X : Il est parfois difficile au premier coup d'œil de bien cerner les aires. Ce type de diagramme est plus efficace lorsqu'il faut comparer un nombre restreint (moins de six) d'éléments, il devient imprécis dès que l'on multiplie le nombre d'entrées. Évitez aussi les inclinaisons en trois dimensions.</p>	<p><a href="#">Évaluation du développement de l'UNICEF (2012-2013)</a> Figures 1 et 2</p>
	<p>Les <b>cartes proportionnelles</b> affichent l'information hiérarchique comme un emboîtement de rectangles dont la taille et la couleur varient en fonction de la valeur associée.</p>	<p>✓ : Un peu comme les diagrammes circulaires, ce type de visualisation montre les parties d'un tout. Il permet donc de voir les tendances et d'effectuer des comparaisons, en plus d'optimiser l'utilisation de l'espace.</p> <p>X : Évitez d'encombrer votre carte proportionnelle avec trop de cases et choisissez judicieusement vos couleurs. Des couleurs trop ternes ou pas assez contrastées rendront le visuel moins efficace, mais faites attention aussi à éviter « l'effet arc-en-ciel ».</p>	<p><a href="#">Atlas of Sustainable Development Goals 2017: From World Development Indicators</a> Figure 1b, p. 2-3.</p>

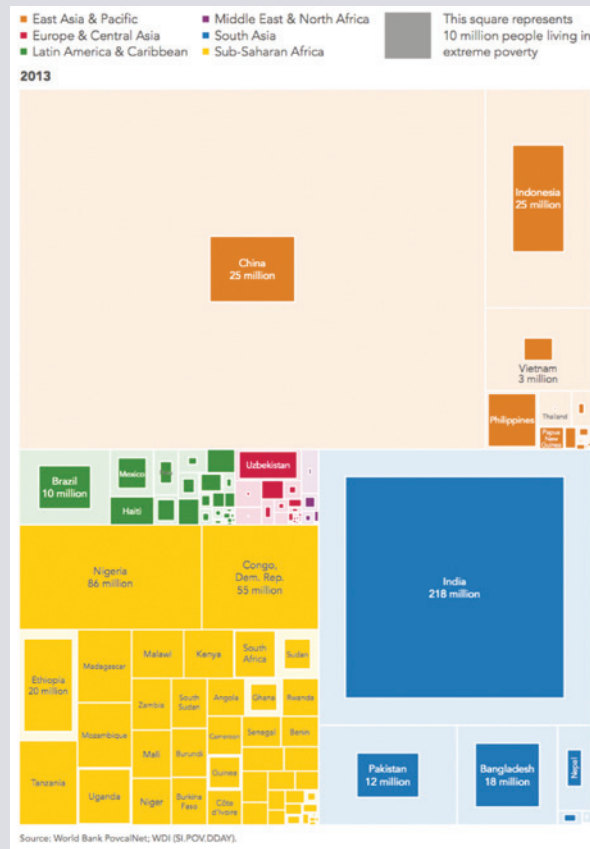


## Encadré 5 – Visualisation des résultats quantitatifs – quelques exemples pour aller plus loin

La visualisation des données permet de fournir un éclairage nouveau et des réponses aux questions qui ont été posées dans le cadre de l'évaluation. Il est courant dans les évaluations de vouloir montrer les performances actuelles vis-à-vis des performances passées, ou encore d'effectuer des comparaisons entre les performances entre différents sites ou groupes. Dans l'exemple ci-dessous, la carte proportionnelle permet de visualiser tous ces éléments rapidement. Le premier diagramme (à gauche) montre le nombre de personnes vivant dans l'extrême pauvreté en 1990 dans différents pays du monde, alors que le deuxième (à droite) démontre l'évolution de la situation plus de 20 ans plus tard (soit en 2013). En voyant cette image, il devient clair que la situation s'est améliorée dans de nombreux pays, par exemple dans les pays asiatiques, d'Amérique latine, du Moyen-Orient et du Nord de l'Afrique, alors que la situation s'est détériorée dans plusieurs pays d'Afrique subsaharienne.



<https://enap.editionsjfd.com>



<https://enap.editionsjfd.com>

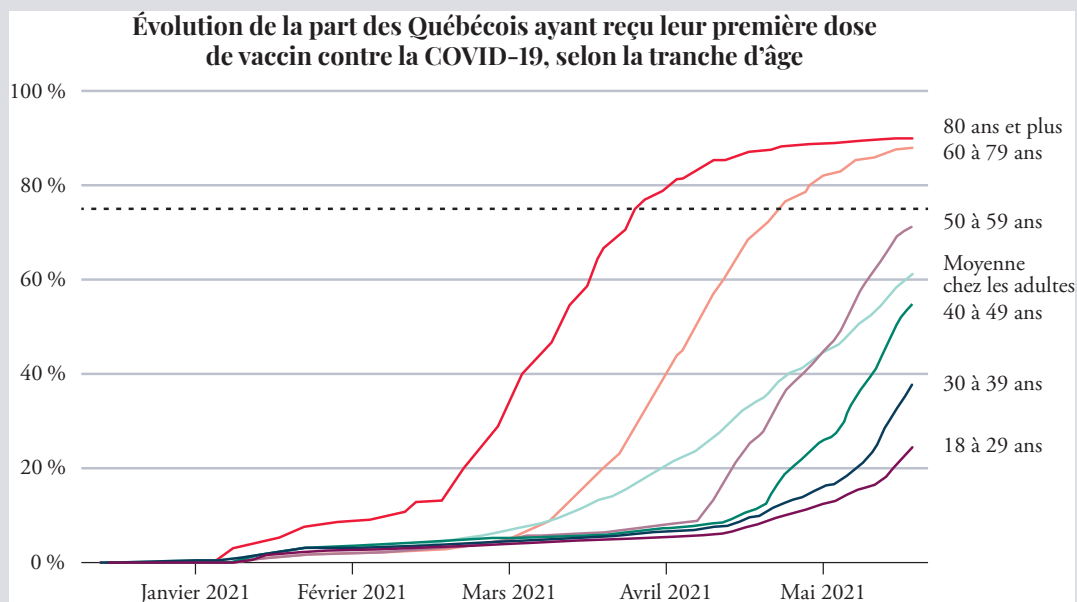
**Figure 2**

Changements dans le nombre de personnes vivant dans l'extrême pauvreté entre 1990 et 2013 dans plusieurs pays du monde.

Source : Reproduit de Banque mondiale, 2017, p. 16-17.

Bien que cette visualisation soit intéressante à plusieurs niveaux, l'utilisation de données agrégées (soit exprimées de manière collective) et de moyennes au niveau national cache de nombreuses réalités particulières. Désagréger (ou ventiler) les données implique de se pencher davantage sur une série de résultats pour souligner les différences qui appartiennent à des sous-groupes particuliers. Il est souvent intéressant de ventiler les données pour faire apparaître une image plus fine de la réalité et pour mettre à jour des tendances qui seraient aussi non cachées. Les données

agrégées peuvent être ventilées par exemple selon le sexe ou le genre, le milieu de vie (urbain/rural/périurbain), les revenus, la langue, l'identité ethnique ou religieuse, l'âge (comme dans l'exemple ci-dessous), etc., pour démontrer les effets différenciés d'une situation sur les individus et rendre les groupes de populations plus vulnérables plus visibles aux concepteurs d'intervention ainsi qu'aux parties intéressées. Dans l'image ci-dessous, qui présente l'évolution de la part des Québécois ayant reçu une première dose de vaccin contre la COVID-19 durant le premier semestre de 2021, on peut voir que la moyenne (données agrégées) ne donne pas une image claire de la réalité, qui présente une grande variabilité selon les groupes d'âge et par rapport à l'objectif à atteindre (démontré par la ligne pointillée dans la portion supérieure du graphique). En désagrégant les données par groupe d'âge, il est possible de faire ressortir les retards et les écarts entre différents groupes de populations, et donc de mieux souligner l'efficacité des services pour la population. En ayant en main cette information, les décideurs sont mieux informés et peuvent prendre de meilleures décisions, par exemple pour mieux cibler leurs actions futures.



**Figure 3**

*Progression de la vaccination contre la COVID-19 au Québec par tranche d'âge, 1<sup>er</sup> semestre de 2021.*

Source : Adapté de LaPresse, à partir des données provenant du ministère de la Santé et des Services Sociaux, cité dans Lacoursière (2021).



Les exemples ci-dessus démontrent aussi que, quel que soit le type de visualisation utilisée, celle-ci devrait :




- être présentée avec un titre clair et complet;
- mentionner les sources des données ainsi que la date de relevé des données;
- avoir une légende ou des étiquettes de données bien identifiées et des axes bien calibrés et clairement définis afin de bien comprendre l'information qui est présentée.



## Moyens qualitatifs de visualisation des résultats

Les résultats qualitatifs permettent souvent de donner une autre perspective, un visage plus humain que les simples chiffres bruts. Quelques façons courantes de visualiser des données qualitatives sont présentées dans le tableau suivant :

**Tableau 3** — Exemples de diagrammes, graphiques et infographies fréquemment utilisés pour visualiser des données qualitatives. Contenu inspiré de : Centre d'assistance Prezi, *Dataviz Project* par Ferdio, Stephanie Evergreen *Qualitative Chart Chooser*, *The Data Visualization Catalogue*. Icônes par : Bhima, Meaghan Hendricks, Product Pencil, Vectplus, Vectors Point, The Noun Project.

Icône	Brève description	Usages appropriés (✓) et à éviter (X)	Lien vers des exemples
	<p>Les <b>citations</b> sont des reproductions d'un court extrait d'un propos ou d'un écrit provenant d'une partie prenante à l'évaluation.</p>	<p>✓ : Les citations peuvent apparaître dans une zone de texte ou dans une barre sur le côté de la page pour les séparer visuellement du reste du texte, être intégrées dans le texte en utilisant des guillemets, ou encore en utilisant des caractères gras afin de les faire ressortir. Les citations sont une façon de donner du poids aux paroles et aux expériences des participants, mais prenez soin d'obtenir et de respecter le consentement des participants, à ce que leurs informations soient incluses dans votre rapport. Pour assurer la confidentialité ou l'anonymat de certaines réponses, vous devrez peut-être attribuer des pseudonymes aux participants, masquer les informations d'identification et/ou présenter des informations globales/agrégées.</p> <p>X : Assurez-vous que les citations correspondent exactement au texte original et clarifiez toujours qu'elles représentent des propos tenus par des participants de l'évaluation.</p>	<p><a href="#">Analyse rapide du genre–COVID-19 Afrique de l'Ouest</a></p> <p>Encadrés p. 3, 7, 9, 10, 13, etc.</p> <p><a href="#">Rapport d'évaluation du programme Passerelle</a></p> <p>Tableau 27, p. 39.</p>
	<p>Une <b>ligne du temps</b> est un moyen graphique d'afficher une liste d'événements en séquence (par exemple, suivant un ordre chronologique).</p>	<p>✓ : La principale fonction des lignes du temps est de communiquer des informations liées au temps, soit pour l'analyse, soit pour présenter visuellement un événement ou un processus, ou encore sa séquence/son déroulement.</p> <p>X : Comme la plupart des visualisations, il faut aller à l'essentiel dans ce qui est présenté afin d'éviter que la visualisation ne devienne complexe et illisible, voire totalement incompréhensible, ce qui en limiterait l'utilité.</p>	<p><a href="#">Rédaction et suivi des recommandations en évaluation–Guide d'accompagnement (gouv.qc.ca)</a></p> <p>Figure 2, p. 6</p> <p><a href="#">The Road to Results: Designing and Conducting Effective Development Evaluations</a>. World Bank</p> <p>Figure 3.3, p. 113</p>

Icône	Brève description	Usages appropriés (✓) et à éviter (X)	Lien vers des exemples
	<p>Les <b>diagrammes de flux</b> montrent des étapes sous forme de cases et indiquent leur ordre en les reliant par des lignes ou des flèches.</p> <p>Ils peuvent être horizontaux, verticaux ou même circulaires.</p>	<p>✓ : Ils sont souvent utilisés pour analyser, concevoir, documenter ou gérer des processus ou des interventions dans divers domaines, des hiérarchies (par exemple, les organigrammes) ou encore des modèles de solutions à un problème donné (modèles logiques ou théories du changement).</p> <p>X : Les phrases longues et compliquées sont à éviter, et chaque case ne devrait contenir qu'un groupe de mots ou des phrases courtes. Ces visualisations étant des modélisations, elles sont souvent réductrices par rapport à la complexité des dynamiques réelles. Bâtir un diagramme de flux qui fasse consensus peut aussi se révéler un processus chronophage. Ils devraient être développés de façon participative pour s'assurer d'une meilleure adhésion des parties prenantes.</p>	<p><a href="#">Rapport sur le développement dans le monde 2017 : Loi et gouvernance</a></p> <p>Figure EO.13.1, p. 30</p>
	<p>Les <b>diagrammes de Sankey</b> sont un type de diagramme de flux dans lequel la largeur des lignes est proportionnelle au flux représenté.</p>	<p>✓ : Ils sont utiles pour visualiser les flux, les processus, les tendances ou même l'évolution de certains éléments, car ils permettent de faire ressortir les éléments majeurs/dominants au sein d'un système complexe.</p> <p>X : Il faut faire attention au surencombrement qui peut facilement rendre la figure illisible.</p>	<p><a href="#">Six Transformations pour atteindre les objectifs de développement durable (ODD)</a></p> <p>Figure 2, p. 10</p> <p><a href="#">Sankey diagram to analyze Relevance   Visual Brains by Sara Vaca</a></p>
	<p>Les <b>graphiques de réseau</b> montrent comment plusieurs éléments sont interconnectés en affichant sous forme de nœuds/ronds des éléments et en les reliant par des lignes qui montrent les relations entre les divers éléments.</p>	<p>✓ : Ce type de graphique est utile pour visualiser des relations complexes entre une grande quantité d'éléments, et aide à mettre en lumière les relations entre les éléments affichés. L'affichage dynamique du réseau peut mettre en évidence des divergences de données qui pourraient autrement passer inaperçues. Des variables supplémentaires peuvent aussi être visualisées, par exemple en faisant en sorte que la taille des nœuds/ronds ou la grosseur des lignes qui les unissent soit proportionnelle à une valeur attribuée.</p> <p>X : Ils peuvent devenir difficiles à lire lorsqu'il y a trop de nœuds et de connexions.</p>	<p><a href="#">Understanding the Role of the World Bank Group in a Crowded Institutional Landscape</a></p> <p><a href="#">IGES Sustainable Development Goals Interlinkages and Network Analysis: A practical tool for SDG integration and policy coherence</a></p> <p>Figure 4, p. 12</p> <p>Figure 5, p. 13</p>

Icône	Brève description	Usages appropriés (✓) et à éviter (X)	Lien vers des exemples
	<p>Les <b>études de cas</b> consistent à utiliser un ou plusieurs exemples réels afin d'obtenir une connaissance approfondie du sujet étudié, et si possible d'en tirer des enseignements pour l'ensemble de l'évaluation.</p>	<p>✓ : Les études de cas permettent de détailler et de mettre en évidence certaines informations importantes. Elles permettent aussi souvent de faire des liens entre la théorie et la pratique. Il est important de bien planifier les études de cas pour qu'elles soient crédibles et qu'elles donnent des résultats satisfaisants.</p> <p>X : La planification, l'analyse des informations et la structuration de l'étude de cas peuvent parfois se révéler complexes, difficiles et chronophages. De plus, les études de cas (et leurs choix) peuvent comporter une certaine partialité.</p>	<p><u>Situation 2020, Vision 2030 : Défis liés aux ODD et solutions pour l'UE</u></p> <p>Témoignages, p. 3, 7, 13, etc.</p> <p><u>Projets pilotes à l'intention des travailleurs âgés : évaluation de la mise en œuvre par études de cas</u></p>
	<p>Les <b>photos</b> sont des images fixes qui reproduisent/représentent des objets réels.</p>	<p>✓ : Les photographies peuvent être utiles à des fins de documentation, pour donner du sens ou pour démontrer certains accomplissements (à des fins de compte rendu, ou en tant que preuves empiriques). Assurez-vous que la signification des dessins ou des photos est clairement expliquée.</p> <p>X : Ce moyen peut comporter une certaine partialité et il faut faire attention à bien respecter les protocoles éthiques, de confidentialité/vie privée et de droits d'auteurs. L'altération ou le trucage des prises de vue est aussi à éviter, tout autant qu'un mauvais échantillonnage de ce qui est présenté, ou encore la présentation de photos de mauvaise qualité (par exemple floues, mal cadrées ou de basse résolution). L'utilisation de nombreuses photos de haute qualité alourdit aussi les fichiers qui peuvent ensuite être difficiles à télécharger.</p>	<p><u>Analyse rapide du genre-COVID-19 Afrique de l'Ouest</u></p> <p>Figure 1, p. 11</p> <p>Figure 2, p. 19</p>

### Encadré 6 – Visualisations des résultats qualitatifs – un exemple pour aller plus loin

Afin de démontrer quelques-uns des résultats atteints dans le cadre d'un projet d'amélioration de la productivité au sein de petites et moyennes entreprises, l'équipe d'évaluation a utilisé des photographies prises par les responsables de la mise en œuvre et par les participants du projet pour documenter les changements au sein de leur environnement de travail. Des échanges avec les participants du projet ont par la suite permis de dégager des tendances dans les expériences vécues par les participants.



**Figure 4**

Documentation des changements apportés aux espaces physiques par le biais de photographies prises par les participants.

Source : iCam Group, 2017

La comparaison des photographies qui démontre le contraste entre la situation initiale et la situation améliorée grâce à l'intervention du projet a aidé les utilisateurs de l'évaluation à comprendre de façon concrète les domaines d'impact du projet (dans ce cas-ci, comment l'ordre et la propreté ont permis des gains d'efficacité dans la gestion des dossiers au sein de l'entreprise). Cet exemple illustre la manière dont la photographie peut être utilisée pour présenter les résultats et démontrer de façon visuelle les changements engendrés par un projet.

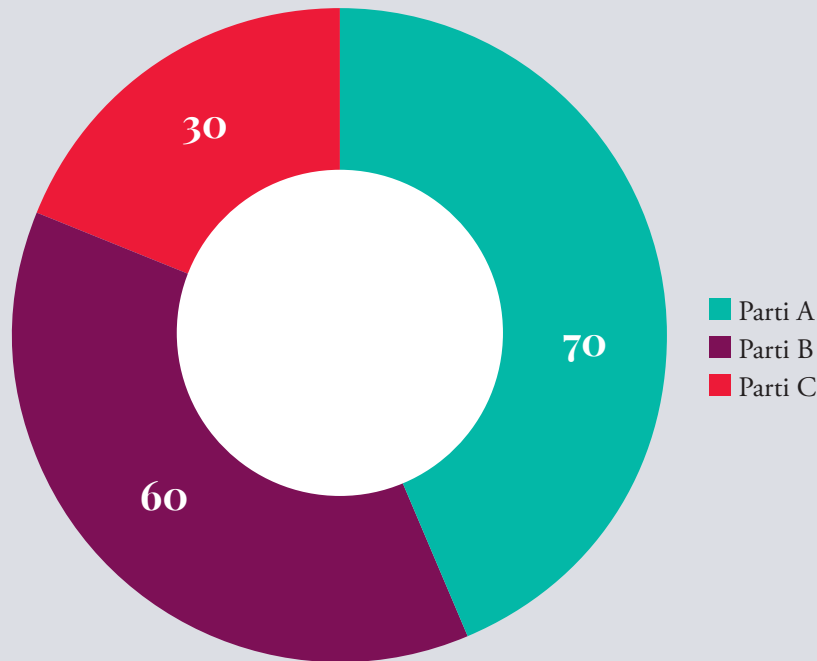
Aussi intéressantes que puissent être ces différentes visualisations des résultats, les visualisations peuvent aussi se révéler nuisibles, notamment lorsqu'elles distraient, déforment ou trompent l'audience. Les visualisations sont à ajouter seulement quand elles sont vraiment nécessaires. Lorsque trop d'informations ou des informations non pertinentes sont fournies, il y a moins de chances que l'information



soit utilisée pour la prise de décision. Il faut faire preuve d'une grande prudence face aux manipulations des données et aux statistiques déformées. Les moyens de visualisation des résultats ne doivent pas servir de moyens d'amplification ou servir à tronquer les résultats.

**Encadré 7 – Des apparences trompeuses. Inspiré de Ayalasomayajula, 2016; Die Zeit, 2017; Calzon, 2021.**

Une mauvaise représentation des données peut fausser leur interprétation. Êtes-vous en mesure de trouver les erreurs dans ces visualisations? Les solutions sont fournies au fur et à mesure.

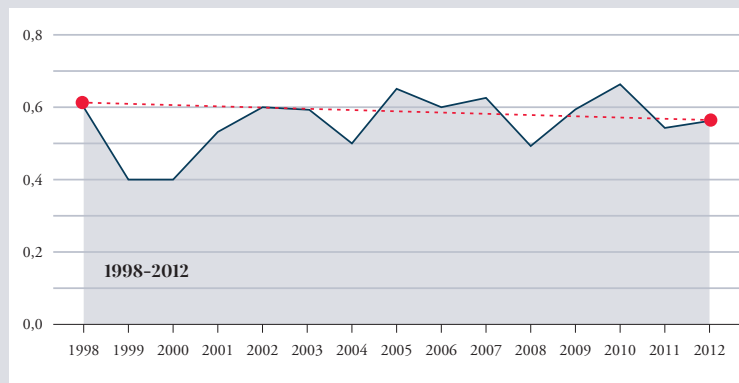


**Figure 5**

Exemple de figure erronée sur les intentions de vote aux prochaines élections en mars 2021, en pourcentage (%).

Les trois secteurs de la tarte totalisent 160 % des intentions de vote, alors que le pourcentage total du cercle devrait représenter 100 %. De telles erreurs dans les données affectent la crédibilité des visualisations. Avant de présenter des données, il est important d'effectuer les contrôles de qualité nécessaires, et ce, à plusieurs étapes (par exemple, lors de la collecte, de la compilation, du nettoyage, des analyses, de la visualisation, etc.).

D'autres erreurs courantes dans les graphiques sont la duplication des données, les données manquantes, les valeurs « non disponibles », non marquées, etc.

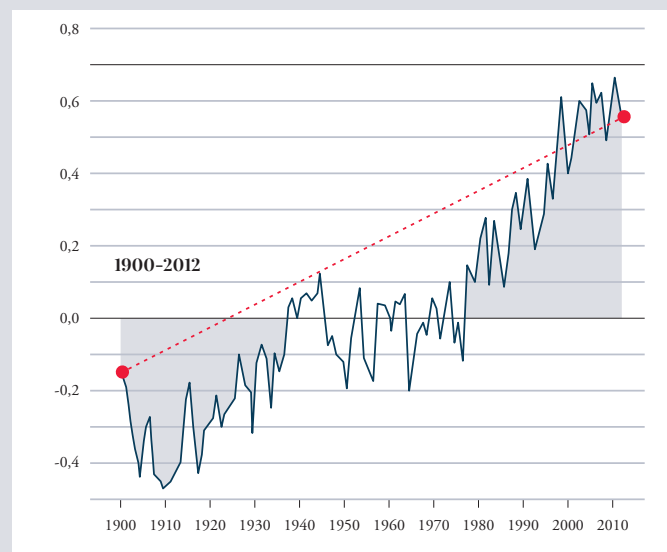


**Figure 6**

Exemple de figure erronée sur l'évolution de la température de l'air à l'échelle mondiale en Celsius, 1998-2012.

Source : adapté de la National Aeronautics and Space Administration (NASA), cité dans Calzon (2021).

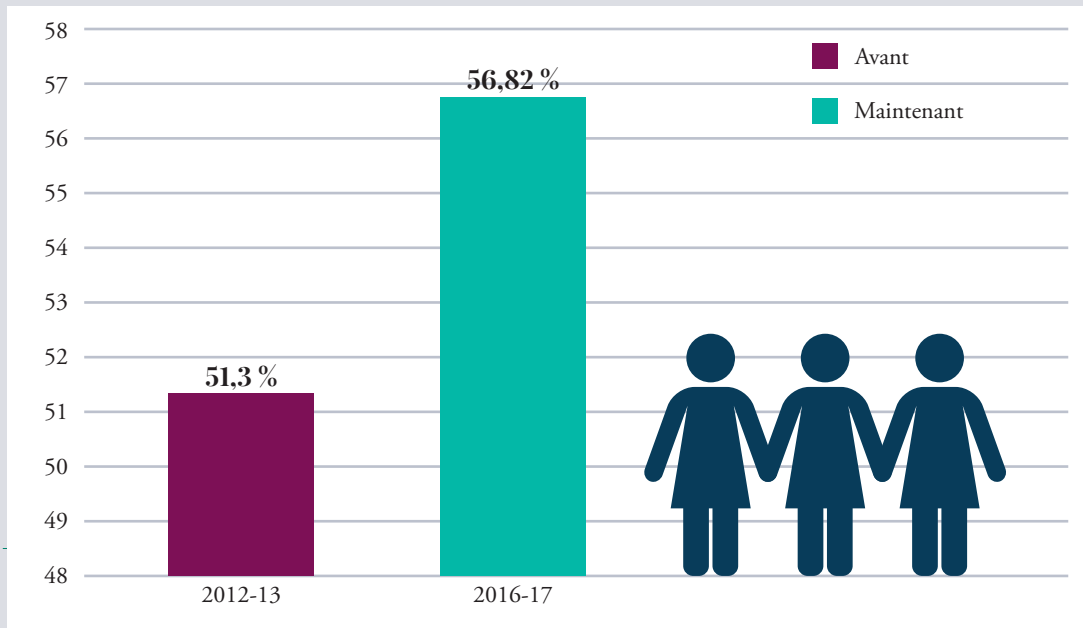
En choisissant de ne montrer qu'une courte période, le graphique fait apparaître les variations comme infimes et peut même faire croire à une tendance plutôt à la baisse! Voyez d'ailleurs la différence dans ce deuxième graphique qui montre les mêmes données, mais présentées sur une période plus longue. Dans ce graphique, plus complet, la tendance au réchauffement est évidente.



**Figure 7**

La bonne façon de présenter l'évolution de la température de l'air à l'échelle mondiale en Celsius, 1900-2012.

Source : adapté de la National Aeronautics and Space Administration (NASA), cité dans Calzon (2021).



**Figure 8**

Exemple de figure erronée sur l'évolution du nombre de femmes bénéficiaires des services, 2012-2017.

Source : adapté de Transforming India, cité dans Ayalasomayajula (2016).

En ne faisant pas commencer l'axe des ordonnées (y) à zéro, le graphique exagère les écarts et les variations, alors qu'en réalité, il y a eu peu de changements (de 51 % à 57 %). Pour représenter correctement ces légères variations, il aurait fallu raccourcir l'ordonnée et les colonnes de manière égale.

## Conclusion

La présentation des résultats constitue un travail exigeant. Lors de la présentation des résultats, les parties prenantes doivent être en mesure de comprendre quels étaient les objectifs de l'évaluation, ce qui a été évalué, comment l'évaluation a été menée, quelles informations ont été obtenues, quelles conclusions ont été tirées et quelles recommandations ont été faites le cas échéant (Joint Committee on Standards, 2004). Un rapport final écrit est une manière reconnue de communiquer et de rendre compte d'une évaluation, mais elle n'est pas la seule.

Pour rendre les rapports écrits, ou toute autre forme de présentation des résultats, plus intéressants, il est possible d'avoir recours à des techniques de visualisation. De bonnes visualisations permettent aux utilisateurs et aux parties prenantes de comprendre rapidement l'information, de repérer les relations entre les éléments, d'examiner les points de données sous un certain angle et de communiquer efficacement grâce aux données. Une bonne visualisation : montre les faits avec précision, attire l'attention de l'utilisateur, complète ou démontre les arguments du texte, comporte un titre et des étiquettes ou bien une légende, est claire et simple (elle n'est pas surchargée), montre les données sans modifier leur message, montre clairement toute tendance ou différence dans les données et doit être exacte en ce qui a trait à l'aspect visuel. Il existe une vaste panoplie de graphiques et autres types de visualisations pour présenter les données. Chacun a un usage spécifique ou qui fonctionne mieux selon le contexte ou le type de données (qualitative ou quantitative) et qu'il est important de garder à l'esprit lorsque vient le temps de décider de la meilleure façon de présenter les résultats.

### Encadré 8 – Questions de réflexion

- Les différents moyens de présentation des résultats ne sont-ils pas complémentaires ?
- En combinant efficacement différents moyens de présentation des résultats (écrit, visualisation, mise en situation, etc.), ne créons-nous pas les conditions nécessaires pour atteindre largement les parties prenantes ?
- En opérant des choix adéquats d'illustration et de représentation visuelle des résultats, n'atteignons-nous pas les hauts standards éthiques et ne gagnons-nous pas en efficacité en matière de communication ?
- En privilégiant la désagrégation des données dans leur traitement et dans la présentation des résultats, ne mettons-nous pas la lumière sur les inégalités de genre et autres, et ne favorisons-nous pas l'émergence de réponses adéquates à ces inégalités ?
- En les présentant de manière efficace et adéquate pour les différentes parties prenantes, ne jetons-nous pas les bases de l'utilisation des résultats ?

## Pour aller plus loin

Cette ressource donne des conseils pour rédiger un rapport d'évaluation. Elle aide à comprendre le rôle et la valeur d'un rapport d'évaluation, à bien comprendre les divers éléments d'un rapport d'évaluation afin de répondre aux besoins d'information des parties prenantes et pour favoriser la participation.

- Centre d'excellence en évaluation. (2004). *Art et méthode de rédaction de rapports d'évaluation*. Secrétariat du Conseil du Trésor du Canada. <https://www.tbs-sct.gc.ca/cee/career-carriere/workshops-ateliers/aawer-amrre-fra.pdf>

Ce document est un module de formation qui porte sur la communication et la reddition de compte efficace d'une évaluation. Il a été produit par *Catholic Relief Services* (CRS) et la Croix-Rouge américaine avec un appui financier des subventions *Food for Peace* (FFP) de l'Agence des États-Unis pour le Développement International (USAID).

- Stetson, V. (2008). *Suivi et évaluation. Communiquer et rendre compte d'une évaluation : directives et outils*. Guy Sharrock (CRS). <http://static1.1.sqspcdn.com/static/f/752898/9984693/1296501611747/www.crs.org>

Ce manuel, produit par Statistique Canada, contient un chapitre entier sur les types de diagrammes et est très axé sur les visualisations des données quantitatives.

Statistique Canada. (2021). *Les statistiques : le pouvoir des données!* <https://www150.statcan.gc.ca/n1/edu/power-pouvoir/toc-tdm/5214718-fra.htm>

Le *Data Viz Project* est un catalogue disponible en ligne qui a pour but d'aider le choix de moyens de visualisation. Pour chaque moyen/visualisation, un descriptif est présenté ainsi que des exemples. Cette ressource est disponible en anglais seulement.

- Ferdio. (s. d.). *Collection of data visualizations to get inspired and finding the right type*. Data Viz Project. <https://datavizproject.com/>

## Références bibliographiques

Alkin, M. (1991). Evaluation theory development. Dans M. W. McLaughlin et D.C. Philips (dir.), *Evaluation and education : At quarter-century* (90<sup>th</sup> yearbook of the National Society for the Study of Education, Part II, p. 91-112). University of Chicago Press.

Ayalasomayajula, V. (2016, 13 septembre). *5 common mistakes that lead to bad data visualization*. Humans of data. <https://humansofdata.atlan.com/2016/09/common-mistakes-bad-data-visualization/#>

Banque mondiale. (2017). *Atlas of Sustainable Development Goals 2017 : From World Development Indicators*. International Bank for Reconstruction and Development / The World Bank. <https://openknowledge.worldbank.org/handle/10986/26306> Licence : CC BY 3.0 IGO

Baron, G. et Monnier, E. (2003). Une approche pluraliste et participative : Coproduire l'évaluation avec la société civile. *Informations sociales*, (110). [https://www.perfeval.pol.ulaval.ca/sites/perfeval.pol.ulaval.ca/files/publication\\_90.pdf](https://www.perfeval.pol.ulaval.ca/sites/perfeval.pol.ulaval.ca/files/publication_90.pdf)

- Beyna, L. (2010). *Performance Monitoring & Evaluation Tips. Constructing an Evaluation Report* (1<sup>re</sup> édition). USAID. [https://pdf.usaid.gov/pdf\\_docs/pnadw117.pdf](https://pdf.usaid.gov/pdf_docs/pnadw117.pdf)
- Calzon, B. (2021, 28 décembre). Misleading Statistics Examples – Discover the potential for misuse of statistics & data in the digital age. *The datapine blog*. <https://www.datapine.com/blog/misleading-statistics-and-data/>
- Caron, D. J. et Bernardi, S. (2019). *La transformation numérique de l'état québécois : un chantier en progression*. École nationale d'administration publique (ENAP). [https://international.enap.ca/international/docs/APlus\\_International/Actualit%C3%A91\\_Gouv%20num\\_ENAP\\_191129\\_mis%20en%20page\\_v2.pdf](https://international.enap.ca/international/docs/APlus_International/Actualit%C3%A91_Gouv%20num_ENAP_191129_mis%20en%20page_v2.pdf)
- Christie, C. A. et Alkin, M. C. (2013). An Evaluation Theory Tree. Dans M. C. Alkin (dir.), *Evaluation roots : A Wider Perspective of Theorists' Views and Influences* (p. 11-58). SAGE Publications.
- Commission économique pour l'Afrique, Union africaine, Groupe de la Banque africaine de développement et Programme des Nations Unies pour le développement. (2016). *OMD-Agenda 2063/ ODD : rapport de transition 2016*. [https://www.afdb.org/fileadmin/uploads/afdb/Documents/Generic-Documents/MDG-SDG\\_Transition\\_Report\\_fre\\_draft3.pdf](https://www.afdb.org/fileadmin/uploads/afdb/Documents/Generic-Documents/MDG-SDG_Transition_Report_fre_draft3.pdf)
- Die Zeit. (2017). Infographie. Illusions graphiques. *Courrier International*. <https://www.courrierinternational.com/grand-format/infographie-illusions-graphiques>
- Duarte, N. (2016). *Slidedocs. Spread ideas with effective visual documents*. Duarte. <https://www.duarte.com/slidedocs/>
- Emery, Y. (2005). La gestion par les résultats dans les organisations publiques : de l'idée aux défis de la réalisation. *Télescope*, automne 2005. [https://serval.unil.ch/resource/serval:BIB\\_38207.P001/REF.pdf](https://serval.unil.ch/resource/serval:BIB_38207.P001/REF.pdf)
- Evergreen, S. (2022). *Qualitative Chart Chooser 3.0*. Evergreendata. Intentional Reporting & Data Visualization [Blog]. <https://stephanieevergreen.com/qualitative-chart-chooser-3/>
- Evergreen, S. (2019). *Effective Data Visualization : The Right Chart for the Right Data*. SAGE Publications.
- Ferdio. (s.d.). *Collection of data visualizations to get inspired and finding the right type*. Data Viz Project. <https://datavizproject.com/>
- IEAG. United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development. (2014). *A World that Counts Mobilising the Data Revolution for Sustainable Development*. <https://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf>
- Jacob, S. et Ouvrard, L. (2009). L'évaluation participative. Avantages et difficultés d'une pratique innovante. *Cahiers de la performance et de l'évaluation*, 1(automne 2009). [http://www.perfeval.pol.ulaval.ca/sites/perfeval.pol.ulaval.ca/files/2009\\_\\_cahiers\\_perfeval\\_1\\_participation.pdf](http://www.perfeval.pol.ulaval.ca/sites/perfeval.pol.ulaval.ca/files/2009__cahiers_perfeval_1_participation.pdf)
- Joint Committee on Standards for Educational Evaluation et Sanders, J. R. (dir.). (1994). *The Program Evaluation Standards : 2nd Edition How to Assess Evaluations of Educational Programs*. SAGE Publications.
- Kallemeyn, L. M. (2018). Expanding the Role of Digital Photographs in Evaluation Practice :

- Documenting, Sense-Making, and Imagining. *Canadian Journal of Program Evaluation / La Revue canadienne d'évaluation de programme*, (33.1 printemps), 114–134. <https://evaluationcanada.ca/fr/system/files/cjpe-entries/33-1-114.pdf>
- King, J. A. (1998). Making sense of participatory evaluation. *New Directions for Evaluation*, (80), 57–67. <https://doi.org/10.1002/ev.1117>
- Kusek Zall, J. et Rist, R. C. (2004). Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners. *World Bank*. <https://openknowledge.worldbank.org/handle/10986/14926>
- Lacoursière, A. (20 mai 2021). Vaccination De cancre à modèle. *La Presse*, <https://www.lapresse.ca/covid-19/2021-05-20/vaccination/de-cancre-a-modele.php>
- Morra Imas, L.G. et Rist, R.C. (2009). *Road to Results : Designing and Conducting Effective Development Evaluations*. The World Bank. <https://openknowledge.worldbank.org/bitstream/handle/10986/2699/52678.pdf?sequence=1&isAllowed=y>
- Patton, M. Q. (2008). *Utilization – focused evaluation*. SAGE Publications.
- Preskill, H. et Torres, R. T. (dir.). (2001). The learning dimension of evaluation use. *New Directions for Evaluation*, 88, 25-37. [https://www.researchgate.net/publication/229747061\\_The\\_learning\\_dimension\\_of\\_evaluation\\_use](https://www.researchgate.net/publication/229747061_The_learning_dimension_of_evaluation_use)
- PNUD IEO. (2018). *Analyse de la qualité des évaluations*. Dans Guide d'évaluation du PNUD (p. 9-12). PNUD IEO. <http://web.undp.org/evaluation/guideline/French/index.shtml>
- Sall, A.A. (2019). *Manuel d'Évaluation des Politiques Publiques*. Section des Comptes de la Cour suprême du Mali.
- Solomon, G.M. Hellstern et H. Wollman (dir.), *Evaluation research and practice: Comparative and international perspectives* (p. 92-106). SAGE Publications.
- Statistique Canada. (2021). *Les statistiques : le pouvoir des données!* <https://www150.statcan.gc.ca/n1/edu/power-pouvoir/toc-tdm/5214718-fra.htm>
- Stetson, V. (2008). *Suivi et évaluation. Communiquer et rendre compte d'une évaluation : directives et outils*. Guy Sharrock (CRS). <http://static1.1.sqspcdn.com/static/f/752898/9984693/1296501611747/www.crs.org>
- Stufflebeam, D. (2001). Interdisciplinary Ph.D. programming in evaluation. *The American Journal of Evaluation*, 22(3), 445-455. <https://journals.sagepub.com/doi/abs/10.1177/109821400102200323?journalCode=ajec>
- USAID. (2010). *Performance Monitoring & Evaluation. TIPS. Constructing an evaluation Report*. [https://pdf.usaid.gov/pdf\\_docs/pnadw117.pdf](https://pdf.usaid.gov/pdf_docs/pnadw117.pdf)
- Voyadzis, C. et Roy, E. (2021). Evaluation of EvalPartner Strategic Plan 2018-2020. *Le Groupe-conseil baastel ltée*. [https://www.evalpartners.org/evidencematters/wp-content/uploads/2021/04/Baastel\\_EP-SP\\_Evaluation-Report-FINAL-2.pdf](https://www.evalpartners.org/evidencematters/wp-content/uploads/2021/04/Baastel_EP-SP_Evaluation-Report-FINAL-2.pdf)
- Wholey, J. (1981). Using evaluation to improve program performance. Dans R.A. Levine, M.A.
- Wholey, J. S. (2001). Defining, improving, and communicating program quality. *Advances in Program Evaluation*, 7, 201–216. [https://doi.org/10.1016/S1474-7863\(01\)80073-9](https://doi.org/10.1016/S1474-7863(01)80073-9)





# L'élaboration des recommandations pour favoriser l'action



*Sylvie Lefrançois  
Serge Éric Yakeu Djiam*

## Introduction

Ce chapitre s'intéresse à différents moyens de présenter les recommandations d'une évaluation afin d'en assurer une bonne réception et une acceptation auprès des différentes parties prenantes en vue de leur utilisation future. Bien que ce ne soit pas toutes les évaluations qui requièrent la présentation de recommandations dans un rapport, lorsqu'elles en exigent, ces dernières doivent être présentées et rédigées afin de favoriser une prise de décision ultérieure. Pour certains mandataires, le terme *recommandation* peut paraître trop « engageant ». On parlera plutôt de suggestions d'amélioration ou de propositions d'amélioration. Quoi qu'il en soit, la plupart des travaux d'évaluation feront état des constats menant à des conclusions et à des pistes d'améliorations futures des différentes interventions évaluées (que ce soient des politiques, des programmes ou des projets). Pour les besoins de ce chapitre, nous utiliserons le terme *recommandation*.

Ce chapitre analyse des moyens permettant de favoriser la réceptivité et l'acceptation des recommandations par les parties prenantes. Il fait également état du rôle de l'évaluateur ou de l'évaluatrice comme intermédiaire entre les parties prenantes et le suivi des recommandations. Des exemples sont ajoutés afin de situer différents contextes de « développement » des pays francophones.

## Diffusion et valorisation : arrimage de l'évaluation avec l'élaboration des politiques publiques

Afin de garantir la réussite du processus de diffusion et de transfert des résultats d'une évaluation, il est important de s'assurer, d'une part, de la réceptivité des mandataires et des principales parties prenantes de l'évaluation et, d'autre part, de leur acceptation des conclusions et des recommandations de celle-ci. Plusieurs éléments seront alors pris en considération. Ceux que nous proposons dans ce chapitre relèvent principalement de la revue de la littérature que nous avons effectuée, mais également de l'expérience que nous avons acquise au fil des années à titre d'évaluateur et d'évaluatrice dans les administrations publiques et les organisations internationales. Les sous-sections ci-dessous analysent les besoins en information, l'identification des parties prenantes et les procédés de formulation des recommandations.

## Considérer les besoins en information

L'usage des recommandations pour l'amélioration de la performance des programmes et des projets ou encore des politiques publiques constitue un élément légitime de la conduite d'une évaluation. Au moment de rédiger les recommandations d'une évaluation, l'évaluateur ou l'évaluatrice doit d'abord passer en revue les principaux résultats et constats énoncés au cours de l'évaluation. La partie IV de cet ouvrage fait état de l'analyse des données, des principaux résultats des évaluations et de la présentation de ces résultats. Comme mentionné, les résultats et constats auront préalablement été présentés et validés par les membres du comité d'évaluation<sup>1</sup>. Ainsi, afin de bien circonscrire les besoins en information, les recommandations sont rédigées de façon à bien répondre aux différents critères<sup>2</sup> couverts par l'évaluation et aux questions d'évaluation en lien avec ces critères.

Les constats d'une évaluation fournissent des évidences tangibles qui soutiennent le bien-fondé des recommandations utiles au processus de prise de décision. Pour assurer leur utilité, il faut que les recommandations soient plausibles et réalisables.

Afin de bien orienter les recommandations, il faut comprendre l'objectif particulier pour lequel l'évaluation est prévue. Rappelons que l'évaluation peut être planifiée dans le but d'élaborer une intervention ou d'aider à sa planification, être réalisée pour déterminer la poursuite, la réorientation ou la réduction d'une intervention, elle peut être menée pour fournir de l'information sur la modification de l'intervention ou simplement, pour contribuer à l'avancement des connaissances. Dans certains cas, une évaluation aura un but tactique pour retarder, par exemple, une décision ou un but politique servant de munition pour un débat politique. L'identification du but principal de l'évaluation orientera également l'évaluateur ou l'évaluatrice dans la formulation des recommandations afin d'en assurer leur valorisation (cf. encadré 1).

### Encadré 1 – Rappel des utilisations possibles des recommandations d'une évaluation

1. Aider à formuler et à justifier les demandes de budget : en mettant l'accent sur les principales conclusions, les recommandations de performance permettent d'orienter les décisions qui peuvent conduire à des augmentations ou à des réductions budgétaires;
2. Aider à motiver le personnel à continuer d'améliorer l'intervention;
3. Formuler et surveiller les performances des organisations et des bénéficiaires;
4. Aider à fournir des services plus efficacement;
5. Soutenir les efforts de planification stratégique et à long terme, fournir des informations de base et suivre ultérieurement les progrès;
6. Permettre de mieux communiquer avec le public cible pour renforcer sa confiance (Kusek et Rist, 2004, p. 139).

1. Vérifier la présentation ultérieure dans l'ouvrage de la description d'un comité de suivi de l'évaluation.

2. Pertinence, efficacité, efficience, impact.

L'utilisation des recommandations peut permettre le développement des connaissances et favoriser l'apprentissage au sein des gouvernements et des organisations. Par exemple, la communauté internationale met de plus en plus l'accent sur l'acquisition des connaissances locales et non sur le transfert des connaissances entre le bailleur de fonds et le mandataire; ce qui favorise davantage les approches d'évaluation basées sur les contextes régionaux et les pratiques locales. Toutefois, l'application de ces informations interpelle les rôles et responsabilités individuelles des parties prenantes impliquées dans le processus d'une évaluation.

Comme cela a été abondamment présenté dans le chapitre sur l'implication des parties prenantes, lors de la réalisation de l'évaluation, ces dernières peuvent jouer un rôle important puisqu'elles ont les connaissances et les perspectives qui peuvent être différentes ou complémentaires à celles de l'équipe d'évaluation. Le même avantage se présente au moment de l'élaboration des recommandations. Ces dernières seront examinées et validées par les parties prenantes. Cette validation permet d'accroître la crédibilité du processus, la transparence et ainsi, l'acceptation des recommandations.

À la fin de la rédaction du rapport d'évaluation, l'évaluateur ou l'évaluatrice doit considérer l'ensemble des questions d'évaluation énoncées dans le mandat et revoir chacun des résultats et constats qui auront été émis afin de valider si toute l'information a pu être recueillie comme prévu, et si des obstacles dans la réalisation du mandat ont pu freiner la validité ou la fiabilité de la collecte. Le lien entre l'assurance des résultats et la capacité de l'évaluation à fournir tous les renseignements nécessaires à chacune des questions d'évaluation permettra d'énoncer des recommandations solides et utiles.

### **Identifier les parties prenantes**

Comme cela a été discuté dans le chapitre 14 : *L'implication des parties prenantes en évaluation*, dans le plan d'évaluation, il est important de bien identifier les rôles et responsabilités des différentes parties prenantes. Fitzpatrick *et al.* (2012) ont élaboré une liste de contrôle pour identifier les parties prenantes. Hurteau *et al.* (2012, p. 128-129) ont adapté cette liste de contrôle pour établir une grille qui prend en considération les différentes personnes requises aux fins de l'évaluation et leurs rôles respectifs. L'identification de ces rôles et responsabilités face à l'intervention permet donc à l'évaluateur ou à l'évaluatrice de distinguer et d'attribuer également des rôles et responsabilités face aux recommandations qu'il ou elle élaborera. Si ces recommandations tiennent compte des différentes capacités d'action des acteurs impliqués, leur réception et leur valorisation seront d'autant plus efficaces et acceptées.

Il faut également prendre en considération le niveau d'intérêt des parties prenantes. Certaines d'entre elles auront des intérêts plus particuliers envers les recommandations émises dans un rapport d'évaluation. La littérature identifie plusieurs façons de déterminer les parties prenantes. Selon le niveau de participation des parties prenantes au processus d'évaluation en fonction de leur intérêt (Agence de la santé et des services sociaux de la Montérégie, s. d.), l'évaluateur ou l'évaluatrice élaborera les recommandations en tenant compte de leur niveau d'importance et d'influence par rapport à l'intervention. Les recommandations seront orientées de façon à accompagner et à guider les parties prenantes qui ont de l'influence et de l'importance par rapport à l'intervention (p. ex. gestionnaires, mandants, bailleurs de fonds) dans la prise de décision ultérieure. Pour être utiles, les recommandations devront également prendre en considération les attentes et les besoins des parties prenantes qui, quant à elles, ont peu d'influence par rapport à l'intervention, mais une grande importance par rapport aux bénéficiaires (p. ex. clientèle cible).

En résumé, les recommandations sont des suggestions pour améliorer, réformer ou renouveler l'intervention évaluée. Elles doivent dériver d'une ou de plusieurs conclusions à partir desquelles ont été identifiés des problèmes spécifiques résumés dans les constats de l'évaluation.

## Formuler des recommandations efficaces

Afin de maximiser l'adhésion des principales parties prenantes à la diffusion et à la valorisation de l'évaluation, l'évaluateur ou l'évaluatrice utilisera une logique d'élaboration des recommandations. Cette logique doit à la fois répondre aux questions de l'évaluation et apprécier la performance de l'intervention, mais surtout, elle sera élaborée de façon que les recommandations soient utiles et présentées en adéquation avec le déroulement de la politique. Plusieurs astuces seront alors utilisées pour favoriser l'utilité des recommandations.

Il est important de noter que les recommandations restent et demeurent des « propositions » d'amélioration des programmes, des projets ou des politiques publiques dans un exercice d'évaluation. Le mandataire a le choix d'accepter ou de rejeter la ou les recommandations qu'il ne juge pas nécessaires ou qu'il souhaite réaliser plus tard.

Bien qu'il n'existe pas de méthode idéale de présentation des recommandations, la formulation de celles-ci devrait être suffisamment détaillée pour permettre aux éventuels décideurs de connaître et de s'approprier les actions à réaliser afin de les appliquer adéquatement et surtout, savoir dans quelle séquence les mettre en application. À notre avis, des recommandations trop générales peuvent être mal interprétées ou encore ne pas être appliquées par les organisations et s'avérer inutiles. Il est possible, au moment de formuler les recommandations, d'entreprendre une discussion avec le mandataire ou le bailleur de fonds de l'évaluation pour vérifier s'il a des souhaits particuliers quant au contenu ou aux spécificités des recommandations qui permettraient de mieux répondre aux besoins de l'évaluation.

Le Secrétariat du Conseil du Trésor du Canada (2004) précise que les recommandations doivent s'appuyer sur des éléments probants. Ainsi :

1. Les recommandations sont étayées par les constatations et les conclusions, dont elles découlent logiquement;
2. Les recommandations portent sur des questions importantes; elles ne constituent pas une liste de souhaits en vrac;
3. Dans la mesure du possible, une analyse des incidences potentielles (sur l'intervention évaluée) de la mise en œuvre de chaque recommandation est présentée;
4. Les recommandations font état des perspectives, des points à améliorer et des possibilités de financement ou de ressources.

On peut tout de même relever deux cas de figure dans la façon de formuler des recommandations dans une évaluation (interne et externe) :

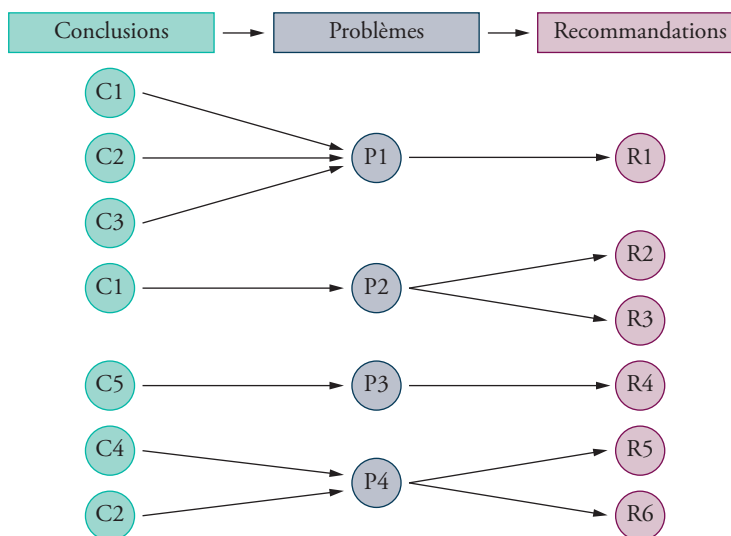
1. **Recommandations élaborées par l'équipe interne d'évaluation** : deux situations peuvent se présenter : i) l'équipe d'évaluation discute des recommandations préparées par le chargé d'évaluation; ii) la réunion de l'équipe d'évaluation est animée de façon à faire émerger des recommandations.

2. **Recommandations « indépendantes » formulées par des évaluateurs ou évaluatrices :** dans ce cas, les recommandations émanent des conclusions et relèvent des principaux constats fiables et vérifiables dans un processus d'évaluation externe. Ici, les échanges entre évaluateurs ou évaluatrices permettent de faire émerger l'adhésion autour des recommandations.

Dans notre pratique professionnelle, nous avons également constaté que des recommandations doivent être énoncées en proposant un échéancier de mise en œuvre de la politique et en fixant des détails sur des indicateurs précis en termes de quantité ou de qualité des mesures. Il ne faut pas oublier que les recommandations, pour être utiles et valorisées, doivent être énoncées de façon pratique, mais également mises en œuvre de façon réaliste. En ce sens, certaines recommandations seront énoncées pour des catégories spécifiques de parties prenantes.

Exemple : Dans ma pratique professionnelle, j'ai adopté depuis plusieurs années une façon bien précise de rédiger des recommandations. Je présente tout d'abord la recommandation dans une phrase clé et je précise par la suite dans un paragraphe les éléments complémentaires, ou quelles seraient les pratiques à adopter pour favoriser l'application de la recommandation. Cela permet d'une part une meilleure compréhension de la recommandation par l'organisation et, d'autre part, d'outiller ou d'orienter l'organisation vers sa mise en œuvre.

La formulation des recommandations est fortement liée aux constats de l'évaluation consolidés sous forme de conclusions préliminaires développées pour chaque critère d'évaluation analysé. Pour une exploitation judicieuse, ces conclusions préliminaires doivent être numérotées ou codifiées afin de recentrer de manière précise les conclusions à partir desquelles sont formulées les recommandations. Les conclusions elles-mêmes peuvent être numérotées et organisées spécifiquement autour des recommandations structurées. Quelle que soit la façon de faire, la figure 1 illustre le processus de formulation des recommandations.



**Figure 1**

Processus de formulation des recommandations.

Il est généralement conseillé de formuler un nombre restreint de recommandations (trois à cinq au maximum pour chaque niveau de recommandations). Au-delà de ce nombre, la pratique a montré soit une confusion, soit une répétition dans la formulation des recommandations.

À titre d'illustration, le cas de la Banque mondiale (*International Finance Corporation* [IFC]) démontré au tableau 2 donne un exemple de matrice simple de formulation et de suivi des recommandations dans une évaluation. Dans cette matrice, les évaluateurs et évaluatrices remplissent les deux premières colonnes, et, par la suite, les gestionnaires de programmes ou de politiques font le suivi de celles-ci et s'assurent qu'elles sont réalisées par leur délégué (Morra Imas et Rist, 2009).

**Tableau 1 — Modèle de formulation et de suivi des recommandations d'une évaluation (IFC).  
Reproduit de Morra Imas et Rist, 2009, p. 474, librement traduit par les auteurs.**

Recommandations	Sources	Date	Responsable	Progrès
1.				
2.				
3.				
4.				
5.				

Dans cet exemple, les recommandations ont une meilleure utilité que si elles sont appliquées. Une façon de donner suite aux recommandations consiste à établir un système de suivi des recommandations. Un tel procédé permet aux parties prenantes d'observer l'évolution de la mise en œuvre des recommandations de l'évaluation. Plusieurs autres modèles de formulation et de suivi des recommandations s'appliquent dans des évaluations en contexte de développement. Les agences du système des Nations Unies en général, et le Fonds des Nations Unies pour l'enfance (UNICEF, 2017 a, 2017b) en particulier, recommandent la formulation des recommandations à deux niveaux : stratégique et opérationnel (cf. tableau 3). Les recommandations stratégiques proposent des solutions au problème posé et identifié sur le long terme, tandis que les recommandations opérationnelles sont des actions à entreprendre sur le court terme pour mettre en œuvre des solutions.



**Tableau 2** — Exemple de modèle de formulation et de suivi des recommandations de l'UNICEF.  
Adapté de UNICEF, 2017b, p. 87.

Numéros des conclusions sur lesquelles repose chaque recommandation	Recommandation stratégique (RS)			
	RS1 : ...			
	Actions clés	Responsables	Ordre de priorité	Commentaires
C1, C2, C...	1. 2.			
C1, C2, C...	<b>RS 2 : ...</b>			
	1. 2.			
C1, C2, C...	<b>RS 3 : ...</b>			
	1. 2.			
<b>Recommandation opérationnelle (RO)</b>				
C1, C2, C...	<b>RO1 : ...</b>			
	1. 2.			
C1, C2, C...	<b>RO2 : ...</b>			
	1. 2.			
C1, C2, C...	<b>RO3 : ...</b>			
	1. 2.			

Comparativement au modèle de l'IFC, celui proposé par l'UNICEF présente un atout dans la mesure où il dispose des propositions d'actions clés et d'un ordre de priorité pour chacune des recommandations formulées. Ainsi, les recommandations sont priorisées et hiérarchisées par l'évaluateur ou l'évaluatrice. En outre, le ou les destinataires de chaque recommandation sont précisés. Soulignons que chaque fois que cela est possible et pertinent suivant le contexte d'application, les recommandations sont élaborées sous forme d'options avec des conditions de mise en œuvre et des conséquences prévisibles de la mise en œuvre pour chacune de ces options.

Enfin, dans un contexte humanitaire, la formulation des recommandations obéit à cette même approche (stratégique et opérationnelle). Ceci tient compte du caractère imprévisible et spontané des activités en cours de réalisation ou à entreprendre. Dans ce cas précis, chaque recommandation proposera des actions et, dans la mesure du possible, des activités spécifiques à développer par l'organisation.

**Tableau 3** — Exemple de présentation des recommandations pour l'évaluation formative de mise en œuvre de la Politique nationale de Protection sociale au Burkina Faso par le Conseil National de la protection sociale (CNPS) avec l'appui technique du Fonds des Nations Unies pour l'enfance (UNICEF) de 2013-20223. Extrait reproduit de UNICEF, 2020, p. 82-84.<sup>3</sup>

Recommandations		
Recommandation stratégique : <i>Adopter une planification des interventions qui renseigne sur les priorités et les régions à couvrir sur une base annuelle en début d'exercice.</i>		
Actions clés	Ordre de priorité	Responsables
<ul style="list-style-type: none"> <li>▪ Revoir le document de la Politique nationale de Protection sociale pour mieux intégrer les défis contextuels liés aux crises humanitaires et relativement aux déplacés internes; cette revue devra également donner plus d'importance aux besoins des personnes vivant avec un handicap;</li> <li>▪ Disposer des référentiels de mise en œuvre utiles à la mesure des performances réalisées pour plus de flexibilité et la mesure des progrès accomplis, et pour une meilleure prise de décision durant les sessions du Conseil National de la protection sociale.</li> </ul>	Court terme	Tous les acteurs
Recommandation opérationnelle : <i>Mettre sur pied et renforcer le dispositif de suivi-évaluation pour de meilleures prises de décisions.</i>		
Actions clés	Ordre de priorité	Responsables
<ul style="list-style-type: none"> <li>▪ Mettre en place un système de suivi et d'évaluation géré par le CNPS avec une responsabilisation des acteurs à tous les niveaux, suivant les prévisions du plan annuel et ses indicateurs de résultats par intervention;</li> <li>▪ Doter ce système de ligne de financement pour un suivi régulier des activités;</li> <li>▪ Adopter un canevas de rapportage unique et harmonisé qui facilite la traçabilité des progrès réalisés par année et par programme de mise en œuvre;</li> <li>▪ Tirer pleinement parti des expériences de suivi des différents acteurs de mise en œuvre.</li> </ul>	Court terme	Conseil National de la protection sociale
Recommandation opérationnelle : <i>Prendre en compte systématiquement le genre.</i>		
Actions clés	Ordre de priorité	Responsables
<ul style="list-style-type: none"> <li>▪ Poursuivre la mise en œuvre en tenant compte du genre;</li> <li>▪ Assurer le suivi en tenant compte du genre par la mise en place d'indicateurs de genre;</li> <li>▪ Élaborer un plan d'intégration systématique du genre ou, lors de la révision de la PNPS, faire une intégration systématique du genre dans le document de l'intervention (cela nécessite une révision du document de catégorisation des bénéficiaires pour désagréger les données par sexe pour toutes les catégories sociales concernées afin de permettre une analyse genre).</li> </ul>	Moyen terme	Secrétariat permanent / Conseil National de la protection sociale

Légende : Court terme = moins de 3 mois

Moyen terme = entre 3 et 6 mois

3. <https://www.unicef.org/evaluation/reports#/detail/15285/evaluation-formative-de-la-mise-en-oeuvre-de-la-politique-nationale-de-protection-sociale-2013-2022-au-burkina-faso> (Consulté le 22 novembre 2021).

## Rôle de l'évaluateur ou évaluatrice comme intermédiaire entre les parties prenantes pour l'élaboration et le suivi des recommandations

### Lors de l'élaboration des recommandations

Au cours d'un processus d'évaluation, une bonne communication avec les principales parties prenantes encourage la connaissance et l'utilisation des résultats et des constats de l'évaluation menant aux recommandations, et permet d'en assurer leur réceptivité. La communication entre l'évaluateur ou l'évaluatrice et les parties prenantes (souvent représentées au sein du comité de suivi de l'évaluation) est en outre indispensable pour garantir une transparence. Une organisation a avantage à faire preuve d'intégrité sur les données probantes de l'évaluation; cela aura pour conséquence de renforcer la crédibilité de l'intervention évaluée et favorisera également la sensibilisation des parties prenantes à l'égard de sa mission.

Une fois les recommandations formulées, il est indispensable de procéder à leur validation. Le processus de validation doit être mutuellement facilité entre l'évaluateur ou l'évaluatrice et le mandataire. L'objectif d'une telle validation est de formaliser l'acceptation des recommandations pour un meilleur suivi. Durant ce processus, les rôles et responsabilités sont identifiés et acceptés par toutes les parties prenantes présentes.

### Lors de la validation des recommandations

Afin d'assurer l'appropriation des recommandations par les parties prenantes, il est suggéré d'organiser des rencontres préliminaires avec les membres du comité de suivi afin de les faire valider. Certains points sont à considérer lors de ces rencontres préliminaires (ex. contraintes budgétaires, ressources humaines, planification, suivi et évaluation). Ainsi, il faut se questionner à savoir si les recommandations sont formulées de façon que chaque élément soit respecté, notamment;

- Le vocabulaire : répond-il adéquatement au style d'expression et au langage utilisé et reconnu par la majorité?
- La compréhension : les parties prenantes ont-elles une compréhension commune du style utilisé dans la présentation des recommandations?
- La réceptivité : les parties prenantes sont-elles en accord avec les recommandations présentées?
- Le réalisme : les recommandations proposées sont-elles en adéquation avec le développement de l'intervention?
- Le nombre : y a-t-il trop de recommandations? Si la réponse est oui, il y a des chances que plusieurs d'entre elles ne soient pas respectées. Il faut s'en tenir aux plus importantes et aux plus réalistes. Il faut que toutes les recommandations soient réalisables, mais ce sont aux parties prenantes de faire les choix et de prendre les décisions sur celles qui seront mises en œuvre.

Ce processus de validation préliminaire a plusieurs avantages. Il permet une grande transparence dans le processus d'évaluation. Il favorise un échange et une discussion sur les meilleures façons possibles de transmettre l'information future afin d'assurer un suivi efficace de l'évaluation et de ses résultats. Il assure une compréhension commune des constats, des conclusions du rapport et de ses recommandations. Finalement, il favorise une plus grande crédibilité du processus d'évaluation.

Pour atteindre les buts poursuivis par l'évaluation, il convient de diffuser les résultats et les conclusions de manière efficace. Pour communiquer efficacement, il faut transmettre des messages clairs, et présenter des données probantes de manière compréhensible et immédiatement accessibles aux parties prenantes. Le chapitre suivant de cet ouvrage exposera plusieurs éléments pour favoriser le transfert des connaissances auprès des populations cibles des interventions évaluées.

## Conclusion

Bien qu'il n'existe pas de procédure idéale pour la rédaction et la formulation des recommandations, nous avons tenté dans ce chapitre de proposer des éléments importants à considérer au cours du processus d'évaluation pour favoriser leur réception et leur appropriation par les parties prenantes. Comme nous l'avons indiqué, une bonne compréhension des différents enjeux et contextes dans lesquels se situe l'évaluation est l'un des points à considérer. La connaissance des principaux rôles et responsabilités des parties prenantes participant au processus d'évaluation est également un élément fort important du processus d'évaluation. Nous avons ajouté à cela quelques techniques utiles et reconnues ainsi que différentes astuces favorisant la rédaction, la présentation, la validation et le suivi des recommandations.

Il ne faut cependant pas oublier l'importance du rôle de l'évaluateur ou de l'évaluatrice dans tout ce processus. En effet, il ou elle doit faire preuve de transparence, d'une capacité d'adaptation et surtout, d'une grande réceptivité envers chacune des parties prenantes. Ces éléments sont gage d'une adhésion des parties prenantes au processus d'évaluation, aux résultats, aux conclusions et aux recommandations de celles-ci. Enfin, tous ces éléments permettront ultimement de favoriser la valorisation de l'évaluation et son apport à la prise de décision.

## Messages clés

Lors de l'élaboration des recommandations d'une évaluation, il faut :

- S'assurer d'avoir bien cerné les besoins de l'évaluation et du contexte dans lequel elle est réalisée;
- Avoir consulté toutes les catégories de parties prenantes permettant de définir tous les éléments importants en lien avec les critères de l'évaluation;
- Élaborer un nombre restreint de recommandations en lien direct avec les principaux constats;
- Rédiger des recommandations constructives permettant à l'organisation d'amorcer le travail rapidement en fonction de leur mise en œuvre.

## Pour aller plus loin

Le lecteur pourra renforcer ses connaissances à travers plusieurs ouvrages. Le livre suivant est vivement recommandé :

Gouvernement du Québec. (2020). *Rédaction et suivi des recommandations en évaluation. Guide d'accompagnement*. [https://www.tresor.gouv.qc.ca/fileadmin/PDF/cadre\\_gestion/Guide\\_redaction\\_suivi\\_recommandations.pdf](https://www.tresor.gouv.qc.ca/fileadmin/PDF/cadre_gestion/Guide_redaction_suivi_recommandations.pdf)

## Références bibliographiques

Agence de la santé et des services sociaux de la Montérégie. (s. d.). *L'analyse des parties prenantes*. [http://extranet.santemonteregie.qc.ca/userfiles/file/performance-innovation/planification-strategique-2010-2015/Approches\\_acteurs\\_2.pdf](http://extranet.santemonteregie.qc.ca/userfiles/file/performance-innovation/planification-strategique-2010-2015/Approches_acteurs_2.pdf)

Agence française de développement et OCDE. (2015). *Évaluation des activités de développement : fonder l'apprentissage et la prise de décision sur des résultats éprouvés*. [https://oecd.org/dac/peer-reviews/12-Lessons-from-the-OECD-DAC\\_FR\\_web.pdf](https://oecd.org/dac/peer-reviews/12-Lessons-from-the-OECD-DAC_FR_web.pdf)

Secrétariat du Conseil du Trésor du Canada. (2004). *Guide pour l'examen des rapports d'évaluation*. <https://www.tbs-sct.gc.ca/cee/tools-outils/grer-gere-fra.pdf>

Fitzpatrick, J. L., Sanders, J. et Worthen, B. (2012). *Program Evaluation : Alternative Approches and Practical Guidelines* (4<sup>e</sup> édition). Pearson Education.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B. et Vermeersch Christel, M. J. (2011). *Evaluation in Practice*. The International Bank for Reconstruction and Development/The World Bank.

Hurteau, M., Bourgeois, I. et Houle, S. (2018). *L'évaluation de programme axée sur la rencontre des acteurs*. Presses de l'Université du Québec.

Hurteau, M., Houle, S. et Guillemette, F. (2012). *L'évaluation de programme axée sur le jugement crédible*. Presses de l'Université du Québec.

Kusek, J. Z. et Rist, R. C. (2004). *Ten Steps to a Results-based Monitoring and Evaluation System : A handbook for development practitioner*. The World Bank.

- Kuzek, J. Z. et Rist, R. C. (2006). *Vers une culture du résultat. Dix étapes pour mettre en place un système de suivi et d'évaluation axé sur les résultats, un guide pour les praticiens du développement*. Les Éditions St-Martin.
- Morra Imas, L. G. et Rist, R. C. (2009). *The Road to Results : Designing and Conducting Effective Development Evaluation*. The World Bank.
- Patton, M. Q. (2014). *Qualitative Research & Evaluation Methods* (4<sup>e</sup> édition). SAGE Publications.
- Snelling, S. et Meserve, A. (2016). *Évaluation des programmes de promotion de la santé : manuel d'introduction*. Santé publique Ontario. <https://www.publichealthontario.ca/-/media/documents/E/2016/evaluating-hp-programs-workbook.pdf?la=fr>
- Stufflebeam, D. L. (1999) *Program Evaluations Metaevaluation Checklist*. Western Michigan University. [https://wmich.edu/sites/default/files/attachments/u350/2014/program\\_metaeval\\_short.pdf](https://wmich.edu/sites/default/files/attachments/u350/2014/program_metaeval_short.pdf)
- UNICEF. (2017a). *UNICEF-Adapted UNEG Evaluation Reports Standards*. <https://www.unicef.org/evaluation/media/816/file/UNICEF-Adapted-UNEG-Evaluation-Report-Standards.pdf>
- UNICEF. (2017b). *Global Evaluation Report Oversight system : Handbook for UNICEF Staff & Independent Assessors*. <https://www.unicef.org/evaluation/media/1381/file/GEROS%20Handbook.pdf>
- UNICEF. (2020). *Évaluation formative de la mise en œuvre de la Politique nationale de Protection sociale 2013 – 2022 au Burkina Faso*. <https://www.unicef.org/evaluation/reports#/detail/15285/evaluation-formative-de-la-mise-en-oeuvre-de-la-politique-nationale-de-protection-sociale-2013-2022-au-burkina-faso>
- Yarbrough, D., Shulha, L. M., Hopson, R. K. et Caruthers, F. A. (2010). *The Program Evaluation Standards : A Guide for Evaluators and Evaluation Users* (3<sup>e</sup> édition). SAGE Publications.

# Un processus de transfert de connaissances pour favoriser l'utilisation des résultats d'évaluation

*Esther Mc Sween-Cadieux*

*Julie Lane*

*Kadidiatou Kadio*

*Valéry Ridde*

*Christian Dagenais*

## Introduction

Les connaissances issues de la recherche et de l'évaluation sont primordiales pour éclairer les politiques, les décisions et les interventions. Toutefois, l'écart persiste, encore et toujours, entre les connaissances disponibles et leur application sur le terrain (Straus *et al.*, 2011). En plus d'occasionner des pertes de temps et d'argent considérables (Meyer, 2010), ce fossé empêche les décideurs, les intervenants et la population de bénéficier des avancées scientifiques (Morris *et al.*, 2011). Certaines études ont d'ailleurs montré que très peu d'innovations ou de découvertes prometteuses étaient concrétisées (Landry et Amara, 2013). Les connaissances disponibles peuvent prendre de 10 à 20 ans pour influencer les pratiques et les décisions (Sussman *et al.*, 2006). Ainsi, bien qu'une quantité phénoménale de connaissances de qualité soient produites chaque année, force est de constater que le processus facilitant le transfert des connaissances (TC) s'avère complexe et pose encore de nombreux défis. De ce fait, l'objectif général de ce chapitre est de vous sensibiliser à l'importance du TC et de vous proposer des outils pour structurer vos démarches de TC afin que vos résultats d'évaluation puissent entraîner des retombées concrètes, et ce, au bénéfice des populations. Le chapitre se décline en deux parties :

- **Partie 1 | Le transfert de connaissances**
  - Qu'est-ce qui peut influencer le transfert de connaissances?
  - Quelle est l'importance d'adopter une approche collaborative en TC?
  - Quel est le potentiel d'être à la fois un évaluateur et un facilitateur du TC?
- **Partie 2 | La planification d'un processus de transfert de connaissances**
  - Quelles sont les étapes clés d'un processus de TC?
  - Comment bien choisir les activités de TC?
  - Comment élaborer un plan de transfert de connaissances?

Tout au long du chapitre, des exemples issus du programme *Équité-Santé* au Burkina Faso ([equitesante.org](http://equitesante.org)) seront utilisés afin d'illustrer les enjeux du transfert de connaissances, mais aussi les façons de bien planifier un processus de TC. Bien que les exemples retenus proviennent du domaine de la santé, les leçons tirées sont tout aussi applicables au TC dans d'autres secteurs. *Équité-Santé* est un programme de recherches interventionnelles issu d'une collaboration entre le Canada, le Burkina Faso et la France. L'objectif visait à comprendre les processus de mise en œuvre d'interventions communautaires pour améliorer l'équité en santé et à en mesurer l'efficacité. Ainsi, certaines interventions déjà en cours dans un district ciblé devaient être évaluées, et d'autres interventions de santé prometteuses devaient être mises en œuvre. Les principales activités du programme se sont intéressées à la lutte contre les maladies fébriles comme le paludisme et la dengue, à la santé maternelle et infantile, aux mutuelles de santé, à la gouvernance du système de santé ainsi qu'aux traumatismes liés aux accidents de la route. Les interventions étudiées ainsi que les questions de recherche ont été développées de manière participative avec les acteurs locaux (p. ex. des professionnels de la santé, des décideurs et gestionnaires, des représentants d'associations et d'organisations non gouvernementales, etc.) lors d'un processus de planification (Dagenais *et al.*, 2015). Ceux-ci, aussi considérés comme des utilisateurs potentiels des connaissances, ont alors pu identifier leurs besoins en connaissances pour améliorer et nourrir leurs pratiques. De plus, un courtier de connaissances (voir encadré) – a été recruté pour mener différentes activités afin de faciliter le dialogue entre les différents acteurs gravitant autour du projet et ainsi, favoriser l'utilisation des connaissances (Dagenais *et al.*, 2016; Mc Sween-Cadieux *et al.*, 2019).

### **Le courtage de connaissances**

*« Le courtage est un processus qui vise, via un acteur intermédiaire, à favoriser les relations et les interactions entre différents acteurs provenant de divers milieux dans le but de produire et d'accompagner l'utilisation des meilleures connaissances disponibles. Les actions de cet intermédiaire peuvent être multiples et variées selon le contexte, mais visent toutes à créer des liens et de la communication, à valoriser les connaissances et les compétences de chacun (sans en prioriser certaines ou sans les hiérarchiser) et à améliorer les pratiques et informer la prise de décision. »* (Définition adaptée de Munerol *et al.*, 2013)

## **Le transfert de connaissances**

En tant qu'évaluateur ou évaluatrice, l'utilité du processus d'évaluation et de ses résultats est au centre de votre pratique afin que les connaissances produites par l'évaluation permettent d'améliorer les programmes que vous évaluez (King et Alkin, 2019). Cependant, le fait de produire des connaissances utiles pouvant mener à des changements n'est bien souvent que le point de départ d'un long chemin si vous souhaitez produire un impact! Effectivement, il faudra déployer de multiples activités pour en favoriser l'utilisation, car ce n'est ni simple, ni linéaire. Le champ du TC s'est développé en réponse aux défis que pose l'utilisation des connaissances. Plus d'une centaine de termes sont utilisés pour identifier ce processus, dont échange, partage, mobilisation, utilisation et en anglais, *knowledge translation*, *knowledge mobilization*, *knowledge exchange*, *knowledge uptake* ou encore, *implementation science* (McKibbin *et al.*, 2010; Esmail *et al.*, 2020) (voir encadré). Dans ce chapitre, nous utiliserons le terme transfert de connaissances (TC), que nous comprenons comme « un processus dynamique et



itératif qui englobe toutes les activités visant la synthèse, la dissémination, l'échange et l'application des connaissances afin de soutenir les décisions et améliorer les pratiques et les politiques dans l'optique d'améliorer la santé des populations» (Instituts de recherche en santé du Canada, 2017). Le TC représente donc un terme parapluie qui englobe les efforts pour mobiliser les connaissances, allant de la synthèse (pour recenser et interpréter les connaissances), à la dissémination (pour informer ou sensibiliser) et à l'échange des connaissances (pour favoriser le partage bidirectionnel des connaissances avec les utilisateurs), jusqu'à l'application des connaissances lorsqu'un changement est requis (ex. : lorsqu'un programme ou une intervention vise à être implanté).

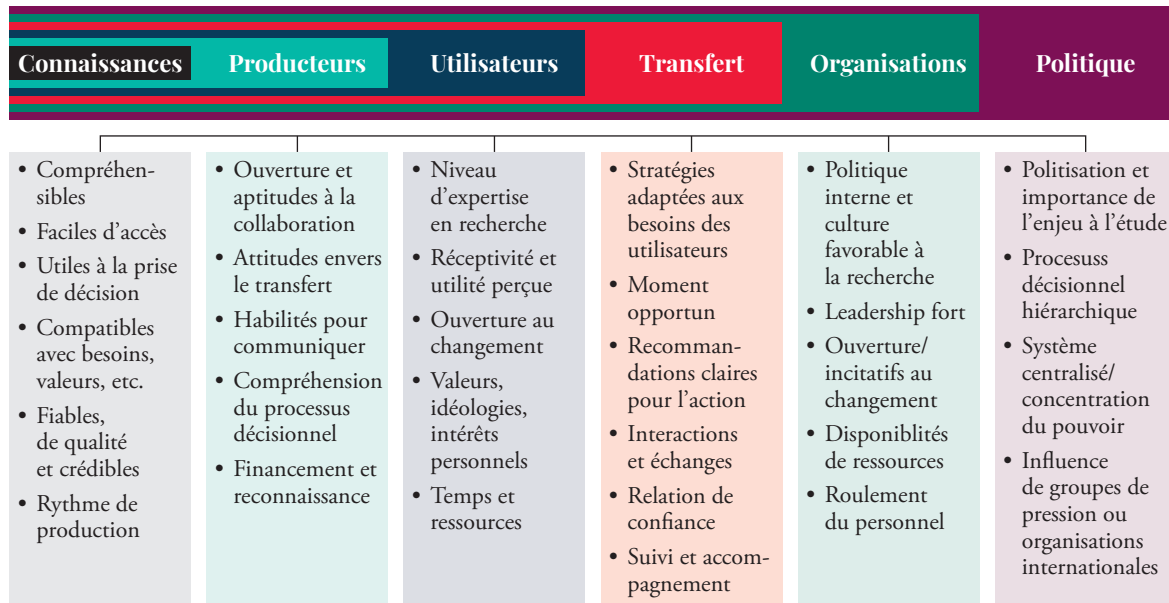
### L'importance de la réflexivité

Bien que le TC ait été initialement conceptualisé comme un processus unidirectionnel allant des producteurs de connaissances scientifiques à des utilisateurs, il est maintenant davantage compris comme un processus misant sur une réciprocité, un échange et un mouvement de connaissances entre les différentes parties prenantes impliquées ou concernées par un enjeu. Toutefois, plusieurs questionnements et débats importants traversent toujours ce domaine de pratique et de recherche puisque les processus peuvent reproduire des dynamiques d'asymétrie de pouvoir, de contrôle, d'oppression et de pratiques colonialistes, autant au niveau de la nature ou des sources des connaissances qui seront considérées et valorisées, que des formes de collaboration ou de la façon dont le partage des connaissances sera réalisé (Stewart, 2019; Smith, 2020; Smylie, 2011). Ainsi, l'inclusion et l'intégration de différents types de connaissances en TC ainsi que la mise en place de processus participatifs qui adressent les inégalités de pouvoir et qui permettent une réelle coproduction des connaissances, sont des principes importants (Boaz *et al.*, 2019). Nous vous invitons donc à réfléchir à ces questions lors de projets de TC en lien avec la décolonisation des savoirs.

Au fil des années, les écrits se sont beaucoup focalisés sur les activités de synthèse, de dissémination et d'échange des connaissances et les facteurs les influençant. Ces écrits ont été plus silencieux sur les activités et processus favorisant l'application des connaissances pour entraîner des changements. C'est en réponse à ce manque que le champ de la science de l'implantation, ou *implementation science*, s'est développé. Ce champ nous offre des enseignements importants sur les processus concrets et les modalités optimales pour qu'une pratique démontrée efficace ou un programme basé sur des données probantes soient adoptés et mis en œuvre en vue d'améliorer la qualité (efficacité, fiabilité, sécurité, pertinence, équité, efficacité) des services offerts à la population (Eccles *et al.*, 2009). La science de l'implantation vise à étudier la «boîte noire» entre la recherche et la pratique pour comprendre comment réussir la mise en œuvre de nouvelles interventions ou programmes en contexte réel, tout en favorisant une évaluation rigoureuse de leurs impacts et retombées (Smith et Polaha, 2017). Il est toutefois important de rappeler que ce ne sont pas toutes les connaissances ou tous les résultats qui peuvent être mis en pratique; certains sont utiles pour sensibiliser à une problématique, améliorer la compréhension ou informer une prise de décision par exemple. Nous reviendrons plus loin sur le fait que le TC peut viser différents objectifs.

## Qu'est-ce qui peut influencer le transfert de connaissances ?

Le TC est un processus multidimensionnel influencé par divers facteurs. Certains peuvent être contrôlés par les évaluateurs ou les responsables de la démarche de TC et d'autres non, car étroitement reliés aux particularités du contexte d'implantation. Il s'avère tout de même primordial d'identifier ces facteurs afin d'adapter nos activités de TC. La figure 1 présente un aperçu des différents obstacles ou facteurs facilitant le TC (Dagenais *et al.*, 2013; Green *et al.*, 2009; Jacobson *et al.*, 2003; Lavis *et al.*, 2003; Oliver *et al.*, 2014; Siron *et al.*, 2015).



**Figure 1**

Les facteurs pouvant faciliter ou entraver le transfert de connaissances

Source : Mc Sween-Cadieux, 2019

Par exemple, il n'est parfois pas suffisant que les connaissances soient pertinentes et utiles à la prise de décision. Si elles ne sont pas compatibles avec les besoins, les valeurs, les attentes des utilisateurs ou avec l'agenda politique, ou encore, si elles ne sont pas perçues comme fiables et de qualité, cela peut nuire à leur utilisation (Orton *et al.*, 2011; Lorenc *et al.*, 2014). Aussi, il est préférable que les résultats arrivent au moment opportun dans la prise de décision pour les parties prenantes, sinon ils risquent de ne pas être utiles (Mitton *et al.*, 2007; Orton *et al.*, 2011). De plus, le rôle des utilisateurs est central, et lorsque ceux-ci s'engagent pleinement dans le processus d'évaluation ou verbalisent leur intention de TC, c'est une condition favorable.

La culture organisationnelle au sein des milieux utilisateurs peut avoir un impact important. Par exemple, la présence d'une politique interne ou d'une culture faisant la promotion du TC, couplée à un leadership fort à l'interne pour la promouvoir, serait un facteur déterminant (Mitton *et al.*, 2007; Nabyonga Orem *et al.*, 2012; Armstrong *et al.*, 2013; Dagenais *et al.*, 2013; Ellen *et al.*, 2014; Budd *et al.*, 2018). Le pouvoir décisionnel des parties prenantes – ou gestionnaires du programme évalué – au sein de l'organisation est également un facteur important (Ellen *et al.*, 2014; Mitton *et al.*, 2007;

Onwujekwe *et al.*, 2015). Par exemple, un système politique très centralisé misant sur une concentration du pouvoir ou un processus décisionnel hiérarchique, peut restreindre les débats et donc, représenter un obstacle au TC (Budd *et al.*, 2018). La prise de décision en lien avec un programme peut aussi subir l'influence des communautés, des débats sociaux, ainsi que l'influence d'autres agences, groupes de pression ou organisations (ex. : les organisations internationales et les bailleurs de fonds dans les pays à faible et moyen revenu) (Nabyonga Orem *et al.*, 2012; Liverani *et al.*, 2013; Budd *et al.*, 2018). Aussi, les limites dans les ressources à disposition pour concrétiser le changement sont également un facteur important.

Finalement, l'aptitude et l'intérêt des évaluateurs à entretenir des liens et à établir des relations égalitaires et collaboratives avec les utilisateurs, voire de les impliquer dans le processus d'évaluation, sont des facteurs importants. Cela peut contribuer à la crédibilité du processus et aplanir certaines résistances possibles face à l'évaluateur (Mitton *et al.*, 2007; Campbell *et al.*, 2009; Hyder *et al.*, 2011). Ainsi, le désir et l'engagement de l'évaluateur à faire en sorte que l'évaluation soit utile et que les résultats soient utilisés sont importants.

### **Encadré 1 – Les facteurs qui influencent le transfert de connaissances**

#### **Un manque de pouvoir décisionnel des districts sanitaires locaux**

Afin de réduire la mortalité maternelle au Burkina Faso, le gouvernement a décidé de subventionner le coût des accouchements pour toutes les femmes à hauteur de 80 % du coût total. Les femmes devaient donc payer 20 %. Cependant, des recherches évaluatives ont montré que ce montant de 20 % constituait toujours une barrière à un accouchement assisté par un professionnel de santé. Pour cela, certaines femmes issues de ménages pauvres et très pauvres sont contraintes à l'accouchement à domicile, augmentant ainsi les risques de complications et de décès maternel. Interpellés par ce résultat et désirant lutter contre l'exclusion des femmes pauvres et très pauvres, les responsables d'un district sanitaire ont décidé de supprimer la contribution de 20 %. Cependant, par souci d'harmonisation de la mise en œuvre de la décision gouvernementale à l'échelle nationale, le ministère de la Santé a ordonné l'arrêt de cette initiative locale. Cet exemple montre que la disponibilité et la pertinence d'un résultat sont insuffisantes pour qu'il soit considéré. Cela montre également que la centralisation du pouvoir de décision au niveau central peut être une barrière au TC.

#### **Un manque de réceptivité et d'utilité perçue par des organisations non gouvernementales (ONG) internationales**

Lors de la planification du programme *Équité-Santé*, un processus d'identification des besoins de recherche et d'évaluation pour améliorer la santé des populations a été mené avec les acteurs locaux d'un district. Ainsi, les représentants locaux d'ONG internationales ont participé à ce processus qui a permis d'identifier des questions d'évaluation pour améliorer leurs interventions communautaires en faveur de meilleurs soins de santé pour les enfants. Les évaluateurs du programme devaient, en collaboration avec les acteurs locaux, évaluer les interventions communautaires et proposer des recommandations. L'application de ces recommandations pouvait donc entraîner une modification des composantes de certaines interventions. Cependant, après deux ou trois rencontres pour mieux clarifier les questions d'évaluation, le courtier de connaissances, qui était l'intermédiaire entre les chercheurs et les représentants des ONG, s'est heurté

à un désintéret et à une méfiance de ces derniers, ce qui a donné lieu à un arrêt de la collaboration. En effet, ces recherches évaluatives n'étaient pas une initiative de la représentation nationale, ni du siège de ces ONG. Les gestionnaires des représentations locales exécutent le plus souvent les directives du niveau national et international. Dès le départ, ce n'était pas un besoin exprimé par ceux qui ont formulé ces interventions communautaires ou qui financent leur mise en œuvre dans le district. Cela illustre que l'intention de l'évaluateur pour produire des recommandations utiles est insuffisante pour susciter la prise en compte de celles-ci. Ainsi, dans cet exemple, le fait que la démarche ne soit pas totalement axée sur les besoins des acteurs qui détiennent un pouvoir de décision a nui à la réussite du TC. Cet exemple montre l'influence que peuvent avoir les attitudes et croyances des utilisateurs envers les résultats d'évaluation (ex. : s'ils sentent que la démarche visait à les contrôler) sur leur utilisation.

### **Quelle est l'importance d'adopter une approche collaborative en TC ?**

Comme vous le devinez peut-être, il est peu probable que vos résultats d'évaluation ou recommandations, diffusés de manière unidirectionnelle dans un rapport adressé à des décideurs, entraînent un changement. Il est nécessaire de faciliter le processus, surtout si les résultats vont à l'encontre de l'ordre déjà en place, des idéologies des acteurs, s'ils nécessitent d'amorcer des changements significatifs ou encore, si le contexte social ou politique n'y est pas propice. À ce sujet, l'évaluateur peut avoir un rôle important à jouer afin de faciliter l'implication des différentes parties prenantes et encourager l'utilisation des résultats du processus (King et Alkin, 2019). Cela est gage d'une plus grande crédibilité du processus pour les utilisateurs et d'une plus grande pertinence des connaissances produites pour leur contexte. À ce sujet, l'implication des parties prenantes dans le processus de TC, grâce à une collaboration égalitaire, est de plus en plus identifiée comme une condition de succès (Jacobson, 2007; Blake et Ottoson, 2009). Par exemple, en misant sur des contacts réguliers pour développer un processus de communication de qualité et une relation de confiance (Mitton *et al.*, 2007; Campbell *et al.*, 2009; Hyder *et al.*, 2011; Oliver *et al.*, 2014; Gagliardi *et al.*, 2016). Les approches interactives en TC misent donc sur la circulation des connaissances de façon bidirectionnelle par la création d'opportunités de rencontres pour clarifier les préoccupations respectives des producteurs et utilisateurs et pour assurer un échange de leurs connaissances et expertises respectives. Par conséquent, les utilisateurs peuvent être impliqués durant les différentes étapes du processus de production et d'appropriation des connaissances (Harris, 2015).

L'importance d'adopter une approche interactive en TC peut aussi être mise en lien avec la place grandissante des approches participatives en évaluation de programme (Ridde, 2006). Cela vient renforcer l'idée que la posture d'expertise – autrefois adoptée par les évaluateurs – doit être délaissée si l'on s'intéresse au TC, et ce, au profit d'une posture d'accompagnement du changement. En adoptant une telle posture, vous pourrez créer un climat de collaboration favorable au partage de vos résultats et de vos recommandations, puisque l'expertise et les connaissances des acteurs sont valorisées et mobilisées dans la démarche de changement.

## Encadré 2 – Importance de favoriser l'implication et l'engagement des acteurs

### L'accompagnement des mutuelles de santé par le courtier de connaissances

Les mutuelles de santé sont des associations locales qui ont pour objectif de favoriser l'accès aux soins de santé de leurs membres par un financement issu de la mise en commun des cotisations des membres. Cependant, le faible taux d'adhésion constitue une difficulté pour leur viabilité. Dans le cadre du programme *Équité-Santé*, les mutuelles de santé d'un district sanitaire ont exprimé le besoin d'un accompagnement pour améliorer le taux d'adhésion. La présence du courtier de connaissances recruté dans le cadre du programme a permis de faciliter ce processus d'accompagnement. Il a d'abord effectué une recension des écrits sur les facteurs de non-adhésion aux mutuelles de santé, qu'il a par la suite synthétisés et rédigés en langage accessible pour les acteurs des mutuelles. Il est ressorti d'une séance de travail que ces facteurs n'étaient pas toujours applicables au contexte local du Burkina Faso. Ainsi, avec l'appui du courtier, les responsables des mutuelles ont mis à profit leurs savoirs et leur expertise du terrain afin d'adapter les constats de la littérature à leurs réalités et de développer et mettre en œuvre un plan d'action afin d'augmenter le nombre d'adhérents à leur mutuelle de santé. Le courtier était en interaction avec les responsables chargés de la gestion des mutuelles, c'est-à-dire que ces derniers détenaient un pouvoir décisionnel concernant l'organisation et le fonctionnement de celles-ci. En plus, les acteurs étaient motivés et disposés à modifier leur pratique. Ce cas illustre l'importance de l'accompagnement offert par le courtier, d'autant plus que l'engagement et la motivation des acteurs et la prise en compte des connaissances qu'ils détiennent peuvent grandement faciliter le TC.

### L'accompagnement des décideurs ministériels par le courtier de connaissances

Toujours dans le cadre du programme *Équité-Santé*, l'un des objectifs du courtier de connaissances recruté était aussi d'appuyer les décideurs du ministère de la Santé dans la prise de décision afin de favoriser la prise en compte des meilleures connaissances. Pour y parvenir, il devait, par exemple, produire des notes synthétiques assorties de recommandations pour orienter les décideurs sur les meilleures stratégies pour résoudre un problème de santé ou pour améliorer les politiques et le système de santé. Il a été proposé que le courtier ait un bureau au sein du ministère de la Santé afin de permettre une plus grande proximité et de faciliter les interactions avec les décideurs. Toutefois, malgré les multiples tentatives, le ministère n'a pas senti le besoin et la nécessité d'accueillir un courtier de connaissances. Ce cas illustre également que l'engagement des utilisateurs potentiels des connaissances est primordial pour permettre à une démarche de TC d'être mise en œuvre.

## Quel est le potentiel d'être à la fois un évaluateur et un facilitateur du TC ?

Plusieurs s'accordent pour dire que ce rôle de facilitation du changement serait central au travail de l'évaluateur (Hartz *et al.*, 2009). Bien que vos recommandations soient utiles et permettent d'améliorer la situation ou de renforcer un programme, les caractéristiques individuelles du messenger et la posture de l'évaluateur dans sa relation avec les utilisateurs influencent le TC. Cette posture de facilitation rappelle celle d'un courtier en connaissances (*knowledge broker*), un rôle d'intermédiation entre l'évaluation et la pratique qui est de plus en plus étudié (Ridde *et al.*, 2013; Bornbaum *et al.*, 2015; Glegg et Hoens, 2016). Le développement du courtage de connaissances s'inscrit dans le

mouvement en TC prônant l'importance de favoriser l'interaction entre les producteurs de connaissances et les utilisateurs (Donnelly *et al.*, 2014). Le but est ainsi de mobiliser les résultats d'évaluation en combinaison avec les connaissances expérientielles des acteurs, et d'en favoriser le TC. Plus précisément, trois grandes catégories d'activités de courtage sont identifiées dans la littérature (Ward *et al.*, 2009; Bornbaum *et al.*, 2015) : 1) la gestion des connaissances (synthèse, adaptation et diffusion des connaissances); 2) la mise en lien avec les acteurs (création d'opportunités d'échanges, développement de réseaux d'acteurs); 3) le renforcement des capacités des acteurs (offrir de la formation et du soutien pour comprendre, évaluer et utiliser les connaissances). Plus récemment, certains auteurs proposent d'élargir ces catégories d'activités en ajoutant : 4) le soutien actif à l'implantation du changement (Albers *et al.*, 2020). Pour les lecteurs qui voudraient en savoir plus à ce sujet, Nilsen (2015) recense les principaux cadres conceptuels, modèles et théories dans le domaine de *l'implementation science*.

Étant à l'interface de plusieurs mondes, un intermédiaire devrait avoir de l'expérience autant dans le milieu de la recherche ou de l'évaluation que dans le milieu d'intervention ou de la prise de décision pour bien comprendre leurs particularités respectives (Mallidou *et al.*, 2018). Il doit démontrer d'excellentes aptitudes à la communication, incluant une facilité à s'exprimer en public ou à modérer des discussions en grand groupe. Bien évidemment, l'utilisation d'un langage clair, simple et efficace pour communiquer avec les différents acteurs est indispensable (Becheikh *et al.*, 2010; Bansal *et al.*, 2012). De solides habiletés interpersonnelles sont aussi importantes afin de développer et surtout maintenir des relations professionnelles avec une multitude d'acteurs ayant des idées, des objectifs et des pouvoirs souvent bien différents (Ward *et al.*, 2009; Phipps et Morton, 2013). Une personnalité entrepreneuriale serait un atout important (Lomas, 2007). Elle se traduit par une capacité de leadership (Mallidou *et al.*, 2018), des habiletés en résolution de problèmes ainsi que des capacités de réseautage, de négociations et de médiation (Phipps et Morton, 2013). Bien souvent, l'intermédiaire doit pouvoir amener différentes parties prenantes à se comprendre mutuellement et à développer un langage commun. La compréhension du processus d'apprentissage des adultes peut également être un atout important, étant donné que l'intermédiaire peut souvent être appelé à offrir des formations et du mentorat aux utilisateurs des connaissances (Glegg et Hoens, 2016).

Par le processus d'évaluation qu'il met en place dans un milieu ou une organisation, un évaluateur peut vraisemblablement gagner à adopter le rôle d'un courtier (ou à s'entourer de collègues pouvant jouer ce rôle!). Le fait de concevoir le processus d'évaluation avec une lunette TC peut l'enrichir en incluant par exemple, d'autres sources de connaissances dans l'évaluation – comme une synthèse des résultats de recherche – ou de créer des liens de collaboration entre le programme évalué et le milieu de la recherche – ou tout autre réseau d'acteurs pertinents (Donnelly *et al.*, 2014). Le processus d'évaluation peut donc être « un important mécanisme pour traduire les connaissances en action » (Donnelly *et al.*, 2014, p. 38). Ainsi, nous croyons important que les compétences spécifiques en TC fassent partie de la formation des évaluateurs, au même titre que les compétences techniques et méthodologiques. L'adoption de ce rôle demande toutefois à l'évaluateur de s'installer dans une posture où il est davantage un facilitateur de processus qu'un expert de contenu. L'expertise de contenu de l'évaluateur demeure, mais il doit la mettre à profit de façon judicieuse dans le processus de TC afin d'éviter de créer des réactions de résistance. Composer avec cette double posture peut être inconfortable pour certains évaluateurs, d'autant plus que ces derniers sont souvent davantage formés à adopter une posture d'expertise. Cette posture rappelle celle proposée par Patton (2010) dans l'approche d'évaluation développementale qui place l'évaluateur davantage comme un accompagnateur du processus de développement de l'intervention dans l'optique de favoriser l'utilisation des résultats.



Toutefois, certains contextes d'évaluation ne sont pas toujours favorables au TC, et ce, indépendamment des efforts de l'évaluateur pour faciliter le processus et pour collaborer avec les parties prenantes, par exemple, lors d'une évaluation sommative « commandée » par les financeurs d'un programme afin de décider s'ils poursuivent ou non le financement. Dans ce cas, une approche de sensibilisation à l'importance du TC et de la plus-value de l'utilisation des résultats d'évaluation pourrait être tentée.

## La planification d'un processus de transfert des connaissances

La première partie de ce chapitre vous a sans doute permis de mieux comprendre en quoi consiste le transfert de connaissances, mais vous vous demandez peut-être par où débiter pour faire en sorte que les résultats de votre évaluation influencent les pratiques, la prise de décisions des décideurs ou l'élaboration des politiques. Quel processus mettre en place ? Comment structurer mes actions ? Dans les prochaines sections, nous mettrons de l'avant l'importance de concevoir vos démarches de TC à partir d'un cadre de référence reconnu et de se doter d'un plan de TC afin de mieux structurer vos démarches visant à faire en sorte que vos résultats d'évaluation puissent entraîner des retombées concrètes. Dans cette partie, nous évoquons les enjeux liés à l'implantation d'un changement en présentant un exemple classique souvent observé dans la lutte contre le paludisme.

### Encadré 3 – Les enjeux liés à l'implantation d'un changement

#### Le cas d'une campagne de sensibilisation dans la lutte contre le paludisme

Le paludisme est une maladie infectieuse propagée par la piqûre de certaines espèces de moustiques. Plusieurs résultats de recherches évaluatives montrent l'importance de dormir sous des moustiquaires imprégnées d'insecticides pour se protéger de ces moustiques. Imaginons une campagne nationale classique de distribution de moustiquaires. Imaginons que la campagne mise sur la distribution de moustiquaires par les municipalités et que ces campagnes, souvent soutenues massivement par le gouvernement du pays, l'Organisation mondiale de la Santé ou des ONG, reposent sur des stratégies de sensibilisation et de communication (ex. : médias, journaux) et une distribution des moustiquaires sur le terrain. De nombreuses campagnes similaires ont eu lieu dans de nombreux pays. Elles se sont soldées par des échecs où plusieurs personnes n'ont pas eu accès à leur moustiquaire, ou d'autres ne l'utilisent pas adéquatement. Pourquoi ? Une telle campagne mise malheureusement sur des activités de diffusion passives réputées inefficaces (ex. : message à la radio). Les organisateurs de ces campagnes, souvent accompagnés de chercheurs spécialisés dans le domaine, bien que bien intentionnés, ne maîtrisent pas les processus et les activités optimales pour implanter de tels changements. Comme dans bien d'autres situations, les activités retenues et organisées ne correspondent pas aux problèmes et enjeux vécus sur le terrain ou aux activités réputées efficaces pour implanter des nouvelles pratiques ou façons de faire. Une analyse des besoins, du contexte, des problèmes et des activités réputées plus efficaces pour entraîner un tel changement de pratique auraient permis de mettre en œuvre un plan de TC susceptible d'engendrer une utilisation plus accrue et durable des moustiquaires par la population.

## Quelles sont les étapes clés d'un processus de transfert ?

Dans le cadre de ce chapitre, nous proposons d'explorer l'un des cadres de référence les plus utilisés, soit le *Knowledge-to-Action Framework* – ou le cadre de la connaissance à l'action (Straus *et al.*, 2013) (figure 2). Ce cadre a été élaboré à la suite de l'analyse d'une soixantaine de cadres de référence et théories (Graham *et al.*, 2006). Il est utilisé par plusieurs programmes et projets internationaux (Curran *et al.*, 2011). Il est particulièrement reconnu pour le fait qu'il est clair; facile d'utilisation; qu'il favorise l'implication des parties prenantes; qu'il prend en considération le contexte local; qu'il propose un processus cyclique et itératif entre les différentes étapes (Peters *et al.*, 2020). Au centre du modèle se trouve «l'entonnoir des connaissances» qui représente le processus par lequel les connaissances sont produites, affinées et adaptées aux besoins des utilisateurs afin qu'elles puissent être appliquées. Une connaissance pourrait référer, par exemple, à une ligne directrice, à une recommandation, à un programme, à une intervention, à une pratique, etc. (Santé Canada, 2017). Autour de «l'entonnoir», le cycle de l'action est un processus qui comporte sept phases nécessaires pour que les connaissances produisent le changement souhaité. Les flèches entre les phases illustrent le fait que celles-ci peuvent s'influencer et que le processus peut être itératif.

→ **La phase 1** vise à identifier les besoins en connaissances des utilisateurs et des publics cibles, ou de cerner un problème qui mérite une attention particulière. Par l'organisation de consultations, elle vise également à identifier les connaissances disponibles sur le sujet, et à analyser l'écart entre ces connaissances et les pratiques actuelles sur le terrain et donc, d'identifier le changement à atteindre.

→ **La phase 2** permet d'adapter les connaissances au contexte local, en collaboration avec les utilisateurs, pour qu'elles deviennent pertinentes et appropriées à leur contexte. L'adaptation doit recourir à un processus transparent, rigoureux et reproductible et idéalement participatif avec les personnes visées par le changement.

→ **La phase 3** permet d'identifier les facteurs qui peuvent freiner ou faciliter l'utilisation des connaissances (p. ex. croyances et ouverture à l'apprentissage, rôles mal définis, manque de leadership, lacunes législatives, etc.). Cette analyse permet de sélectionner des activités de TC pour répondre aux obstacles identifiés et de tirer profit des facteurs favorables.

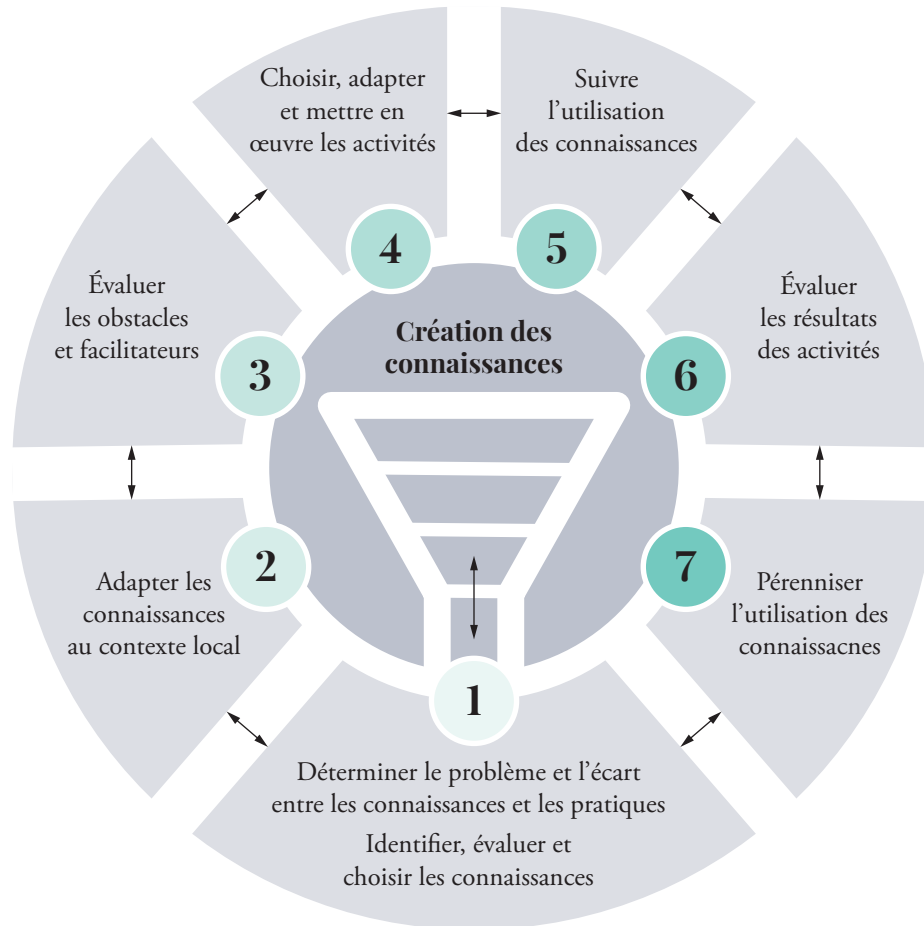
→ **La phase 4** permet de choisir et d'adapter les activités en fonction des facteurs précédemment identifiés, du contexte de mise en œuvre et de celles qui sont réputées plus efficaces pour faciliter le TC. Cette étape vise également l'identification des acteurs spécifiques à impliquer dans le processus d'implantation et même, de créer des équipes ou des comités pour accompagner la mise en œuvre.

→ **La phase 5** vise à suivre l'utilisation des connaissances et les changements au niveau des connaissances, de la compréhension, des attitudes, de la modification des comportements, etc. Il est également recommandé de suivre la qualité de la mise en œuvre des activités afin de s'assurer que les résultats attendus pourront être produits.

→ **La phase 6** vise à évaluer les retombées avec une méthodologie explicite et rigoureuse, qualitative et/ou quantitative, pour une évaluation au niveau structurel (ex. : culture organisationnelle, changements dans le système), du processus (ex. : satisfaction des professionnels envers les pratiques professionnelles) ou des résultats ultimes (ex. : bien-être des populations).



→ **La phase 7** permet de vérifier, par des mécanismes de suivi ou de soutien, que les changements se maintiennent dans le temps. Cette pérennité doit être considérée dès le début du processus de TC en tenant compte des ressources budgétaires et humaines disponibles, mais aussi du leadership, de la nature des partenariats, des capacités organisationnelles, etc.



**Figure 2**

Le cadre de référence « de la connaissance à l'action ».

Source : adapté de Straus, Tetroe et Graham, 2013; traduction libre par les auteurs

### Comment bien choisir les activités de TC ?

Une limite de ce cadre de référence est que la phase 4 n'explique pas la façon de choisir les activités de TC les plus appropriées et la façon de structurer un plan de transfert de connaissances. Trop souvent, les évaluateurs choisissent des activités sans s'assurer qu'elles concordent avec leur cible d'action. Par exemple, certains utilisent des activités qui permettent davantage la sensibilisation de leurs publics cibles à leurs résultats (ex. : la diffusion d'informations sur un site Internet) alors que leur cible d'action réelle est de vouloir influencer une politique publique qui est en élaboration. Le choix des activités à mettre en œuvre va donc varier selon l'objectif et les retombées visées, par exemple sensibiliser ou susciter l'intérêt, influencer ou vaincre des préjugés, améliorer l'action grâce à un changement de

pratique chez les professionnels, influencer une prise de décision politique ou mobiliser l'action publique, etc. Ce choix devra aussi être adapté aux besoins des publics cibles, au contexte de mise en œuvre, mais aussi à la nature et au type de connaissances à partager (Graham et Tetroe, 2009; Barwick *et al.*, 2014). Aussi, certains font l'hypothèse que plus l'objectif du TC est ambitieux, comme changer les pratiques ou informer une décision politique, plus les efforts devront être importants et soutenus dans le temps (Landry *et al.*, 2001) et inclure des échanges interactifs et fréquents (Lemire *et al.*, 2009) et des activités de soutien à l'implantation (ex. : *coaching*, mentorat, formation, courtage de connaissances, communauté de pratique, etc.).

Pour vous guider dans le choix des activités, la taxonomie développée par Walter *et al.* (2005) (Nutley *et al.*, 2007) est utile. Contrairement à d'autres classifications qui s'appliquent davantage aux changements de pratiques des professionnels de la santé, cette taxonomie est intersectorielle. Elle est basée autant sur les écrits en sciences sociales, en santé qu'en éducation. Cinq mécanismes de changements sont proposés pour catégoriser les activités visant le TC :

→ **Dissémination** : les activités qui permettent de diffuser les connaissances de manière adaptées (c'est-à-dire sous forme de messages clés) aux publics cibles (ex. : note de politique, fiche synthèse, matériel éducatif, guides de pratique, infographie, vidéo, ateliers et séminaires, etc.).

→ **Interaction** : les activités qui permettent de renforcer les liens de collaboration entre, par exemple, les praticiens, les décideurs et les chercheurs ou évaluateurs dans le but de combiner les connaissances scientifiques et expérientielles (ex. : démarche partenariale ou coconstruction, implication des utilisateurs dans le processus d'évaluation, comité consultatif, processus délibératif).

→ **Influence sociale** : les activités qui misent sur l'influence d'individus significatifs (ex. : leaders d'opinion, experts locaux, champions, courtage de connaissances) pour sensibiliser les utilisateurs à l'importance des connaissances, modifier leurs attitudes et croyances et agir sur le plan de la motivation au changement.

→ **Facilitation** : les activités qui visent à faciliter le transfert et l'implantation des connaissances en offrant un soutien autant technique, financier, organisationnel qu'émotionnel. Elles permettent de créer un contexte propice au changement en fournissant l'assistance et les moyens nécessaires à l'action (ex. : formations personnalisées, mentorat, outils technologiques, communauté de pratique).

→ **Incitatifs et renforcement** : les activités qui misent sur l'utilisation de récompenses ou d'autres stratégies d'encouragement et de motivation pour renforcer l'utilisation des connaissances (ex. : audit et rétroaction, système de rappels, incitatifs financiers, changement de statut professionnel, etc.).

Le prochain encadré reprend l'exemple de la campagne de sensibilisation à l'utilisation des moustiquaires dans la lutte contre le paludisme en utilisant le cadre *Knowledge-to-Action* pour guider le processus visant l'implantation d'un changement par la population. Vous pourrez constater que la stratégie de départ, misant principalement sur des activités de dissémination, gagnerait grandement à miser sur d'autres mécanismes comme celui de l'influence sociale.

#### Encadré 4 – Illustration du cadre de référence *Knowledge-to-action*

En s'inspirant du cadre conceptuel de référence *Knowledge-to-Action* et des mécanismes de changement, les évaluateurs impliqués dans de telles campagnes pourraient sensibiliser les organisateurs d'une campagne nationale classique de distribution de moustiquaires imprégnées d'insecticides à longue durée d'action (MILDA) pour contrer le paludisme à l'importance de :

- Faire une analyse de la situation et des besoins dans les familles afin de voir l'utilisation ou la non-utilisation actuelle des MILDA;
- Analyser la nécessité d'adapter les MILDA aux besoins de certaines familles (ex. : MILDA de plus grand format pour les familles avec un seul lit);
- Identifier les facteurs qui facilitent ou qui freinent l'utilisation des MILDA (ex. : certaines familles utilisent le MILDA pour d'autres finalités);
- Sélectionner et adapter les activités de dissémination (ex. : le message radio) et des activités d'influence sociale, par exemple nommer des acteurs communautaires dans chacun des villages, qui seront crédibles aux yeux des villageois et qui seront préalablement formés à jouer leur rôle. Ils pourront accompagner les familles à réaliser la pertinence des MILDA, les guider à l'utilisation quotidienne des MILDA (les aider à comprendre le pourquoi, le comment, le quand), à identifier des activités incitatives et de renforcement pour contrer les freins à l'utilisation des MILDA, renforcer l'habitude des familles dans l'utilisation des MILDA, etc.;
- Suivre l'utilisation des connaissances en invitant les champions dans chacun des villages à faire le tour des familles de façon périodique afin de vérifier si les familles utilisent toujours les MILDA, et accompagner les familles qui ne les utilisent plus;
- Évaluer les effets par un suivi d'indicateurs (ex. : nombre de familles qui utilisent les MILDA, taux de familles qui consultent à la clinique pour la malaria);
- Favoriser la pérennité de l'utilisation des connaissances en mettant en place différentes stratégies visant à rappeler aux familles l'importance d'utiliser les MILDA.

#### Comment élaborer un plan de transfert des connaissances ?

Bien que la créativité et l'expérience pratique soient nécessaires pour choisir les activités qui ont le plus de chances d'atteindre l'objectif de transfert de connaissances, il s'avère primordial d'adopter une approche structurée et d'explicitier la théorie du changement – ou logique d'action – sous-jacente à la démarche de TC. C'est pourquoi, après avoir survolé les activités potentielles pour favoriser le TC, nous vous proposons d'élaborer un plan de TC (Barwick, 2008; Lemire *et al.*, 2009). Le plan de TC peut être imbriqué dans un plan d'évaluation et ne constitue donc pas une démarche à mener en parallèle. Pour vous guider dans votre réflexion, voici huit questions complémentaires importantes à considérer.

**Quoi transférer?** Vous devez ici préciser le message que vous souhaitez transférer. Dans le flot de données et de résultats d'évaluation que vous avez en main, quels sont ceux que vous souhaitez transférer? Il importe de préciser vos messages, car trop souvent, nous avons tendance à les noyer dans une mer d'informations où nous sommes portés à ne pas être vraiment clairs avec les messages que nous souhaitons transférer.

**Quel est l'objectif?** Vous devez ici préciser les objectifs que vous visez. Souhaitez-vous, par exemple, informer et faciliter la compréhension d'un enjeu, sensibiliser et favoriser un changement d'attitude, mobiliser l'action et susciter l'engagement, soutenir et éclairer une prise de décision, améliorer et accompagner des changements de pratiques professionnelles ou de comportements auprès de la population ciblée? En fonction des acteurs que vous souhaitez rejoindre, de leurs besoins ainsi que de la nature des connaissances, précisez les objectifs que vous visez ainsi que l'horizon temporel pour les atteindre.

**À qui transférer?** Vous devez préciser ici le ou les publics cibles ou les futurs utilisateurs des connaissances. Qui aurait le pouvoir d'agir sur les recommandations ou sur les messages transmis pour atteindre l'objectif souhaité? Voulez-vous rejoindre des chercheurs, des décideurs, des intervenants ou la population en général? Peut-être souhaitez-vous rejoindre l'ensemble de ces acteurs. Si tel est le cas, vous devrez adapter vos messages aux caractéristiques de chacun de ces acteurs.

**Qui devrait transférer?** Vous devez ici préciser qui sera le messager. Réfléchissez bien ici au profil de votre messager en fonction des activités que vous allez vouloir mettre en œuvre. Par exemple, vous souhaitez influencer la prise de décision de décideurs. Imaginons que ces derniers sont à réviser une politique et que vous souhaitez les influencer en portant à leur attention certains résultats d'évaluation. Vous leur proposez un atelier délibératif. Vous devrez vous assurer que votre messager (qui sera l'animateur de cet atelier) sera un interlocuteur crédible aux yeux des décideurs.

**Comment transférer?** Vous devez ici préciser les activités que vous allez privilégier pour atteindre vos objectifs. En fonction des objectifs de TC ciblés, vous devez préciser les activités ou modalités de TC pour atteindre ces objectifs en vous inspirant, par exemple, de la taxonomie développée par Walter *et al.* (2005) et présentée aux pages 16 et 17. Ainsi, si votre objectif principal est d'informer le public cible sur un enjeu, des activités de dissémination pourraient être pertinentes (p. ex. rédaction d'une note de politique et organisation d'un atelier avec les parties prenantes). Ne pas oublier que si l'objectif est ambitieux et vise un changement, il serait préférable que les activités soient interactives, soutenues dans le temps et ancrées dans un processus collaboratif solide.

**Quand transférer?** Vous devez préciser ici les moments où vous allez vouloir mettre en œuvre les activités de transfert des connaissances, en tenant compte du calendrier et des besoins des utilisateurs afin de saisir les moments propices au TC. À ce sujet, trop souvent, les porteurs d'un projet attendent à la fin pour planifier et réaliser des activités de TC, alors qu'il serait bénéfique de s'y attarder dès le début. Il est important que certaines activités vous permettent d'être en lien avec les utilisateurs afin d'aborder la question du TC et de l'utilisation des résultats, de déterminer conjointement des activités de TC et d'actualiser le plan en cours de projet afin de vous assurer qu'il demeure pertinent et qu'il répond aux besoins des utilisateurs. Cela peut se concrétiser par la mise sur pied d'un comité consultatif ou de suivi du TC.

**Comment évaluer les effets?** Vous devez préciser ici comment vous allez évaluer si le transfert des connaissances a été réussi, c'est-à-dire que les objectifs fixés ont été atteints à court, moyen et long terme. Vous devrez préciser les indicateurs que vous allez vouloir évaluer (p. ex. les perceptions ou les comportements de certains intervenants) ainsi que les méthodologies pour les évaluer (p. ex. passation d'un questionnaire après les activités ou réalisation d'entretiens).

**Combien de ressources?** Vous devez préciser ici les ressources (financières, humaines et autres) qui seront nécessaires pour actualiser votre plan de transfert des connaissances et son évaluation. Trop souvent, les porteurs de projets ne budgètent pas cette partie et arrivent à la fin d'un projet sans ressources pour le transfert des connaissances.

### Encadré 5 – Le plan de transfert de connaissances

#### Des activités de TC pour favoriser l'adhésion aux mutuelles de santé au Burkina Faso\*.

Dans le cadre du programme *Équité-Santé*, un courtier de connaissances a été recruté afin de favoriser le TC auprès des partenaires locaux d'un district sanitaire. Lors de l'atelier de planification du programme en 2012, un groupe de responsables de mutuelles de santé a exprimé le besoin d'être soutenu dans la réalisation et l'amélioration de ses activités par le courtier de connaissances. Ces responsables souhaitaient comprendre les déterminants de la faible adhésion des populations à leur mutuelle de santé. Pour illustrer les différentes composantes du plan de TC, nous utilisons le travail d'accompagnement que le courtier a réalisé auprès de ce groupe d'utilisateurs.

**Quoi transférer?** Le courtier a effectué une recension des écrits sur les déterminants de la faible adhésion des populations à leur mutuelle de santé afin de synthétiser et d'interpréter les connaissances les plus pertinentes. Le paiement des cotisations à la mutuelle en deux tranches plutôt qu'un paiement unique, l'importance de la sensibilisation de la population au bien-fondé des mutuelles et l'implication des leaders communautaires (chefs des villages) dans les mutuelles sont des exemples de recommandations proposées par le courtier.

**Quel est l'objectif du TC?** Améliorer les pratiques des mutuelles de santé (résultat intermédiaire), en vue d'améliorer la couverture de leur mutuelle (résultat ultime). Les responsables avaient déjà exprimé leur ouverture à modifier le fonctionnement de leurs mutuelles.

**À qui transférer?** Les responsables de mutuelles de santé du district partenaire du programme *Équité-Santé* représentent les principaux utilisateurs des connaissances.

**Qui devrait transférer?** Le courtier de connaissances est crédible aux yeux des utilisateurs des connaissances. De plus, il a développé des compétences en courtage de connaissances dans le cadre d'une formation spécialisée.

**Comment transférer?** Le courtier a organisé un atelier délibératif en octobre 2012 avec les responsables des mutuelles. Ensuite, il a rédigé une note de politiques dans laquelle il énonce des recommandations claires aux mutualistes pour favoriser l'adhésion de la population. Le courtier a également préparé avec les responsables un plan d'action adapté au contexte burkinabé concernant la mise en œuvre des recommandations. Des rencontres de suivi entre le courtier

et les responsables des mutuelles ont aussi été prévues tous les trois mois pour poursuivre l'accompagnement. Il s'avère bien souvent nécessaire de combiner plusieurs activités de TC pour favoriser l'atteinte des objectifs.

**Quand transférer?** Des rencontres de suivi trimestrielles ont été prévues pour accompagner la mise en œuvre du plan d'action. La fréquence des différentes activités de TC prévues a été établie en concertation avec les responsables des mutuelles en fonction de leurs préférences et de leur disponibilité pour y participer.

**Comment évaluer les effets?** Afin de s'assurer de la pertinence des activités du courtier, les mutualistes seront invités à répondre à des questionnaires de satisfaction à la fin des ateliers et des réunions de suivi. Des entretiens individuels ont également été prévus pour évaluer leur satisfaction ainsi que les changements dans leurs pratiques depuis le début de l'accompagnement par le courtier. Les effets des activités de courtage seront également évalués à l'aide, par exemple, de l'augmentation du nombre d'adhérents aux mutuelles à la suite de la mise en œuvre de leur plan d'action.

**Combien de ressources?** Les frais liés à l'organisation des ateliers et des réunions de suivi devaient être planifiés (ex. : frais de déplacement, frais de graphisme et d'impression des notes de politiques et autres documents, location de salle, etc.). Les frais liés aux heures de travail du courtier étaient également prévus dans le programme *Équité-Santé*.

\* Ce cas illustre un exemple de TC qui nécessite des ressources (p. ex. acteur intermédiaire) qui ne sont pas à la portée de tous. Plusieurs démarches de TC peuvent toutefois nécessiter moins de ressources ou utiliser des ressources existantes. Par exemple, il est possible d'intégrer certaines activités de TC aux responsabilités d'une personne du projet, de combiner des activités de TC à des rencontres déjà planifiées à d'autres fins ou encore, d'utiliser des indicateurs de suivi des données qui ont déjà été collectées, etc.

## Conclusion

Nous espérons vous avoir sensibilisé à l'importance d'adopter une lunette *transfert de connaissances* dans vos projets d'évaluation, mais aussi de vous avoir outillé à la planification d'une démarche de transfert et d'implantation des connaissances. Puisque ce chapitre s'adressait avant tout à des évaluateurs, nous croyons important de conclure en abordant l'importance de mettre à profit vos compétences, et d'évaluer vos démarches de TC. Malgré la multiplication des termes et cadres conceptuels et l'intérêt grandissant pour ce domaine de recherche, d'importants besoins demeurent pour mieux comprendre les mécanismes de TC les plus efficaces et pour développer des méthodologies adaptées à l'évaluation du TC.

## Messages clés

- Les objectifs et finalités possibles d'une démarche de TC sont multiples et le processus peut être influencé par de multiples facteurs liés par exemple aux connaissances à transférer, aux individus impliqués dans la démarche, aux activités de transfert de connaissances, à la culture organisationnelle et aux particularités contextuelles.
- Les évaluateurs gagneraient à adopter davantage le rôle d'un courtier de connaissances pour favoriser le TC en développant diverses compétences techniques, mais aussi relationnelles, pour l'animation de processus et de soutien à l'implantation.
- L'utilisation du cadre de référence *Knowledge-to-Action* et l'élaboration d'un plan de transfert de connaissances permettront de mieux structurer vos démarches visant à faire en sorte que vos résultats d'évaluation puissent entraîner des retombées concrètes (p. ex. influencer les pratiques, la prise de décision, l'élaboration de politiques, etc.).
- Le choix de vos activités de TC sera modulé en fonction de votre objectif et des retombées ou type d'utilisation que vous visez. Ce choix doit être adapté aux besoins et aux préférences des publics cibles, au contexte de mise en œuvre, mais aussi à la nature et au type de connaissances à partager.

## Pour aller plus loin

Si vous souhaitez en connaître davantage sur le domaine du transfert de connaissances, voici trois ressources en français, disponibles gratuitement en ligne :

→ **Une série de formations en ligne (de type MOOC ou cours en ligne ouverts à tous (CLOT))** élaborées par l'Équipe RENARD en collaboration avec l'Institut de recherche pour le développement (IRD) et du Centre de pédagogie universitaire (CPU) de l'Université de Montréal. Le premier cours est une introduction au TC, le deuxième porte sur la note de politique et le troisième s'intéresse au courtage de connaissances.

<https://catalogue.edulib.org/fr/series/le-transfert-de-connaissances/>

→ **La trousse d'accompagnement SACO** pour l'implantation et l'évaluation des stratégies de transfert et d'utilisation des connaissances. Cette trousse répertorie les principales activités de TC et pour chacune, vous trouverez toute l'information nécessaire sur ce qu'est cette activité, quand l'utiliser, quelle est son efficacité, quel est le processus d'implantation et plus encore.

<http://www.saco.uqam.ca/>

→ **Le référentiel de compétences en transfert de connaissances** est un outil présentant les connaissances et les comportements observables requis pour mener à bien les tâches et les responsabilités liées aux diverses activités propres à la pratique du transfert de connaissances.

<https://referentieltc.ca/>



## Références bibliographiques

- Albers, B., Metz, A. et Burke, K. (2020). Implementation support practitioners – a proposal for consolidating a diverse evidence base. *BMC Health Services Research*, 20(1), 368. <https://doi.org/10.1186/s12913-020-05145-1>
- Armstrong, R., Waters, E., Dobbins, M., Anderson, L., Moore, L., Petticrew, M., Clarck, R., Pettman, T. L., Burns, C., Moodie, M., Conning, R. et Swinburn, B. (2013). Knowledge translation strategies to improve the use of evidence in public health decision making in local government : intervention design and implementation plan. *Implementation Science*, 8, 121. <https://doi.org/10.1186/1748-5908-8-121>
- Armstrong, R., Waters, E., Moore, L., Dobbins, M., Pettman, T., Burns, C., Swinburn, B., Anderson, L. et Petticrew, M. (2014). Understanding evidence : a statewide survey to explore evidence-informed public health decision-making in a local government setting. *Implementation Science*, 9(1), 188. <https://doi.org/10.1186/s13012-014-0188-7>
- Bansal, P., Bertels, S., Ewart, T., MacConnachie, P. et O'Brien, J. (2012). Bridging the Research-Practice Gap. *Academy of Management Perspectives*, 26(1), 73–92. <https://doi.org/10.5465/amp.2011.0140>
- Barwick, M. (2008). *Knowledge Translation Planning Template*. Ontario : The Hospital for Sick Children.
- Barwick, M., Coriandoli, R., Johnny, M., Myers, G. et Phipps, D. (2014). Knowledge Translation and Strategic Communications : Unpacking Differences and Similarities for Scholarly and Research Communications. *Scholarly and Research Communication*, 5(3), 1-14.
- Becheikh, N., Castonguay, Y., Landry, R., Fakhreddine, M. O. I. et Ziam, S. (2010). How to improve knowledge transfer strategies and practices in education? Answers from a systematic literature review. *Research in higher education journal*, 7, 1–21.
- Blake, S. C. and Ottoson, J. M. (2009). Knowledge Utilization : Implications for Evaluation. *New Directions for Evaluation*, (124), 21–34.
- Boaz, A., Davies, H., Fraser, A. et Nutley, S. (dir.). (2019). *What works now? Evidence-informed policy and practice*. Policy Press.
- Bornbaum, C. C., Kornas, K., Peirson, L. et Rosella, L. C. (2015). Exploring the function and effectiveness of knowledge brokers as facilitators of knowledge translation in health-related settings : a systematic review and thematic analysis. *Implementation Science*, 10, 162. <https://doi.org/10.1186/s13012-015-0351-9>
- Budd, E. L., Armstrong, R., Becker, L., Brownson, R. C., deRuyter, A. J., Furtado, K. S., Mui, T., Pettman, T., Reis, R. S., Saunders, T., Shi, J., Sung-Chan, P., Wang, Z. et Ying, X. (2018). A qualitative exploration of contextual factors that influence dissemination and implementation of evidence-based chronic disease prevention across four countries. *BMC Health Services Research*, 18(1), 233. <https://doi.org/10.1186/s12913-018-3054-5>
- Campbell, D. M., Cooke, M., Jorm, L., Redman, S., Rychetnik, L. et Zwi, A. B. (2009). Increasing the use of evidence in health policy : practice and views of policy makers and researchers. *Australia and New Zealand Health Policy*, 6(1), 21. <https://doi.org/10.1186/1743-8462-6-21>

- Curran, J. A., Campbell, B., Grimshaw, J. et Hayden, J. A. (2011). Knowledge translation research : The science of moving research into policy and practice. *Journal of Continuing Education in the Health Professions*, 31(3), 174–180. <https://doi.org/10.1002/chp.20124>
- Dagenais, C., Boileau-Falardeau, M., McSween-Cadieux, E., Ridde, V. et Somé, T. D. (2015). Collaborative development and implementation of a knowledge brokering program to promote research use in Burkina Faso, West Africa. *Global Health Action*, 8. <https://doi.org/10.3402/gha.v8.26004>
- Dagenais, C., McSween-Cadieux, E., Ridde, V. et Somé, P.-A. (2016). A Knowledge Brokering Program in Burkina Faso (West Africa) : Reflections from Our Experience. *Health Systems & Reform*, 2(4), 367–372. <https://doi.org/10.1080/23288604.2016.1202368>
- Dagenais, C., Queuille, L. et Ridde, V. (2013). Evaluation of a knowledge transfer strategy from a user fee exemption program for vulnerable populations in Burkina Faso. *Global Health Promotion*, 20(1 Suppl.), 70–79. <https://doi.org/10.1177/1757975912462416>
- Donnelly, C., Klinger, D., Letts, L. J. et Shulha, L. (2014). Supporting Knowledge Translation Through Evaluation : Evaluator as Knowledge Broker. *Canadian Journal of Program Evaluation*, 29(1).
- Eccles, M. P., Armstrong, D., Baker, R., Cleary, K., Davies, H., Davies, S., Glasziou, P., Iltis, I., Kinmonth, A.-L., Leng, G., Logan, S., Marteau, T., Michie, S., Rogers, H., Rycroft-Malone, J. et Sibbald, B. (2009). An implementation research agenda. *Implementation science*, 4(18). <https://doi.org/10.1186/1748-5908-4-18>
- Ellen, M. E., Bouchard, G., Grimshaw, J. M., Lavis, J. N., Léon, G. et Ouimet, M. (2014). Barriers, facilitators and views about next steps to implementing supports for evidence-informed decision-making in health systems : a qualitative study. *Implementation Science*, 9(1), 179. <https://doi.org/10.1186/s13012-014-0179-8>
- Esmail, R., Brown, S., Clement, F. M., Hanson, H. M., Holroyd-Leduc, J., Niven, D. J., Straus, S. E. et Striffler, L. (2020). A scoping review of full-spectrum knowledge translation theories, models, and frameworks. *Implementation Science*, 15(1), 1-14. <https://doi.org/10.1186/s13012-020-0964-5>
- Gagliardi, A. R., Berta, W., Boyko, J., Kothari, A. et Urquhart, R. (2016). Integrated knowledge translation (IKT) in health care : a scoping review. *Implementation Science*, 11(38). <https://doi.org/10.1186/s13012-016-0399-1>
- Glegg, S. M. et Hoens, A. (2016). Role Domains of Knowledge Brokering : A Model for the Health Care Setting. *Journal of neurologic physical therapy*, 40(2), 115–123. <http://dx.doi.org/10.1097/NPT.000000000000122>
- Graham, I. D., Logan, J., Harrison, M., Straus, S., Tetroe, J., Caswell, W. et Robinson, N. (2006). Lost in knowledge translation : time for a map? *The Journal of Continuing Education in the Health Professions*, 26(1), 13–24. <https://doi.org/10.1002/chp.47>
- Graham, I. D. et Tetroe, J. M. (2009). Getting Evidence into Policy and Practice : Perspective of a Health Research Funder. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 18(1), 46–50.
- Green, L. W., García, C., Hiatt, R. A. et Ottoson, J. M. (2009). Diffusion Theory and Knowledge Dissemination, Utilization, and Integration in Public Health. *Annual Review of Public Health*, 30(1), 151–174. <https://doi.org/10.1146/annurev.publhealth.031308.100049>

- Harris, R. (2015). The Impact of Research on Development Policy and Practice : This Much We Know. Dans A. Chib, J. May et R. Barrantes (dir.), *Impact of Information Society Research in the Global South* (p. 21-43). Springer.
- Hartz, Z. M. A., Denis, J.-L., Matida, A. et Moreira, E. (2009). From Knowledge to Action : Challenges and Opportunities for Increasing the Use of Evaluation in Health Promotion Policies and Practices. Dans L. Potvin, L. M. Anderson, L. de Salazar, M. Hall, Z. M. A. Hartz et D. V. McQueen (dir.), *Health Promotion Evaluation Practices in the Americas : Values and Research* (p. 101-120). Springer.
- Hyder, A. A., Corluka, A., El-Shinnawy, A., Ghaffar, A., Ghassany, H., Lim, M.-K., Malekafzali, H., Mfutso-Bengo, J., Segura, E. et Winch, P. J. (2011). National policy-makers speak out : are researchers giving them what they need? *Health Policy and Planning*, 26(1), 73–82. <https://doi.org/10.1093/heapol/czq020>
- Instituts de recherche en santé du Canada (IRSC). (2017). *Application des connaissances*. (Consulté le 18 août 2020). <http://www.cihir-irsc.gc.ca/f/29529.html>
- Jacobson, N. (2007). Social epistemology : Theory for the “fourth wave” of knowledge transfer and exchange research. *Science Communication*, 29(1), 116–127. <https://doi.org/10.1177/1075547007305166>
- Jacobson, N., Butterill, D. et Goering, P. (2003). Development of a framework for knowledge translation : understanding user context. *Journal of Health Services Research & Policy*, 8(2), 94–99. <https://doi.org/10.1258/10755470321466067>
- King, J. A. et Alkin, M. C. (2019). The Centrality of Use : Theories of Evaluation Use and Influence and Thoughts on the First 50 Years of Use Research. *American Journal of Evaluation*. SAGE Publications Inc., 40(3), 431–458. <https://doi.org/10.1177/10755470198214018796328>
- Landry, R. et Amara, N. (2013). From knowledge transfer to knowledge management and value creation. Dans L. Potvin et P. Armstrong (dir.), *Shaping Academia for the Public Good : Critical Reflections on the CHSRF/CIHR Chair Program* (p. 51-78). University of Toronto Press.
- Landry, R., Amara, N. et Lamari, M. (2001). Climbing the Ladder of Research Utilization : Evidence from Social Science Research. *Science Communication*, 22(4), 396–422. <http://dx.doi.org/10.1177/1075547001022004003>
- Lavis, J. N., Abelson, J., McLeod, C. B., Robertson, D. et Woodside, J. M. (2003). How can research organizations more effectively transfer research knowledge to decision makers? *The Milbank Quarterly*, 81(2), 221–248. <https://doi.org/10.1111/1468-0009.t01-1-00052>
- Lemire, N., Souffez, K. et Laurendeau, M.-C. (2009). *Animer un processus de transfert des connaissances : bilan des connaissances et outil d'animation*. Institut national de santé publique du Québec.
- Liverani, M., Hawkins, B. et Parkhurst, J. O. (2013). Political and Institutional Influences on the Use of Evidence in Public Health Policy. A Systematic Review. *PLOS ONE*, 8(10), e77404. <https://doi.org/10.1371/journal.pone.0077404>
- Lomas, J. (2007). The in-between world of knowledge brokering. *BMJ : British Medical Journal*, 334(7585), 129.
- Lorenc, T., Duffy, S., Lock, K., Martineau, F. P., Petticrew, M., Phillips, G. et Tyner, E. F. (2014). Cultures of evidence across policy sectors : systematic review of qualitative evidence. *European Journal of Public Health*, 24(6), 1041–1047. <http://dx.doi.org/10.1093/eurpub/cku038>

- Mallidou, A. A., Atherton, P., Chan, L., Frisch, N., Glegg, S. et Scarrow, G. (2018). Core knowledge translation competencies : a scoping review. *BMC Health Services Research*, 18(1), 502. <https://doi.org/10.1186/s12913-018-3314-4>
- Mc Sween-Cadieux, E. (2019). *Évaluation de stratégies de transfert de connaissances mises en œuvre dans le cadre du programme Équité-Santé au Burkina Faso* [thèse de doctorat, Université de Montréal]. <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/22686>
- Mc Sween-Cadieux, E., Dagenais, C., Ridde, V. et Somé, D. T. (2019). A health knowledge brokering intervention in a district of Burkina Faso : A qualitative retrospective implementation analysis. *PLOS ONE*, 14(7), e0220105. <https://doi.org/10.1371/journal.pone.0220105>
- McKibbon, K. A., Ciliska, D., Davis, D. A., Dobbins, M., Haynes, R. B., Lokker, C., Straus, S. E. et Wilczynski, N. L. (2010). A cross-sectional study of the number and frequency of terms used to refer to knowledge translation in a body of health literature in 2006 : a Tower of Babel? *Implementation Science*, 5(1), 16. doi : 10.1186/1748-5908-5-16
- Meyer, M. (2010). The Rise of the Knowledge Broker. *Science Communication*, 32(1), 118–127. <https://doi.org/10.1177/1075547009359797>
- Mitton, C., Adair, C. E., McKenzie, E., Patten, S. B. et Perry, B. W. (2007). Knowledge Transfer and Exchange : Review and Synthesis of the Literature. *The Milbank Quarterly*, 85(4), 729–768. <https://doi.org/10.1111/j.1468-0009.2007.00506.x>
- Morris, Z. S., Wooding, S. et Grant, J. (2011). The answer is 17 years, what is the question : understanding time lags in translational research. *Journal of the Royal Society of Medicine*, 104(12), 510–520. <https://doi.org/10.1258/jrsm.2011.110180>
- Munerol, L., Cambon, L. et Alla, F. (2013). Le courtage en connaissances, définition et mise en œuvre : une revue de la littérature. *Santé publique*, 25(5), 587-597.
- Nabyonga Orem, J., Mafigiri, D. K., Marchal, B., Ssenooba, F., Macq, J. et Criel, B. (2012). Research, evidence and policymaking : the perspectives of policy actors on improving uptake of evidence in health policy development and implementation in Uganda. *BMC Public Health*, 12(1), 109. <https://doi.org/10.1186/1471-2458-12-109>
- Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science*, 10, 53. <https://doi.org/10.1186/s13012-015-0242-0>
- Nutley, S. M., Walter, I. et Davies, H. T. O. (2007). *Using evidence : How research can inform public services*. Policy Press, University of Bristol.
- Oliver, K., Innvar, S., Lorenc, T., Thomas, J. et Woodman, J. (2014). A systematic review of barriers to and facilitators of the use of evidence by policymakers. *BMC Health Services Research*, 14(1), 2. <https://doi.org/10.1186/1472-6963-14-2>
- Onwujekwe, O., Etiaba, E., Mbachu, C., Mirzoev, T., Russo, G., Uguru, N. et Uzochukwu, B. (2015). Role and use of evidence in policymaking : an analysis of case studies from the health sector in Nigeria. *Health Research Policy and Systems*, 13(1), 46. <https://doi.org/10.1186/s12961-015-0049-0>
- Orton, L., Capewell, S., Lloyd-Williams, F., O’Flaherty, M. et Taylor-Robinson, D. (2011). The Use of Research Evidence in Public Health Decision Making Processes : Systematic Review. *PLOS ONE*, 6(7), e21704. <https://doi.org/10.1371/journal.pone.0021704>

- Patton, M. Q. (2010). *Developmental Evaluation : Applying Complexity Concepts to Enhance Innovation and Use*. Guilford Press.
- Peters, S., Bussi eres, A., Cristens, J., Depreitere, B., Thomas, A., Vanholle, S. et Vermandere, M. (2020). Facilitating Guideline Implementation in Primary Health Care Practices. *Journal of Primary Care & Community Health*, 11. <https://doi.org/10.1177%2F2150132720916263>
- Phipps, D. et Morton, S. (2013). Qualities of knowledge brokers : reflections from practice. *Evidence & Policy*, 9(2), 255–265. <https://doi.org/10.1332/174426413X667784>
- Ridde, V. (2006). Suggestions d'am eliorations d'un cadre conceptuel de l' valuation participative. *The Canadian Journal of Program Evaluation*, 21(2), 1–23.
- Ridde, V., Dagenais, C. et Boileau-Falardeau, M. (2013). Une synth ese exploratoire du courtage en connaissance en sant  publique. *Sant  Publique*, 25(2), p. 137–145.
- Sant  Canada. (2017). *Guide de planification de l'application des connaissances*. [https://publications.gc.ca/collections/collection\\_2018/sc-hc/H14-226-2017-fra.pdf](https://publications.gc.ca/collections/collection_2018/sc-hc/H14-226-2017-fra.pdf)
- Siron, S., Dagenais, C. et Ridde, V. (2015). What research tells us about knowledge transfer strategies to improve public health in low-income countries : a scoping review. *International Journal of Public Health*, 60(7), 849–863. <https://doi.org/10.1007/s00038-015-0716-5>
- Smith, H. C. (2020). *Understanding the role of community in knowledge mobilization* [th ese de doctorat, Universit  York]. YorkS [https://yorkspace.library.yorku.ca/xmlui/bitstream/handle/10315/37738/Smith\\_Hilda\\_C\\_2020\\_PhD.pdf?sequence=2&isAllowed=y](https://yorkspace.library.yorku.ca/xmlui/bitstream/handle/10315/37738/Smith_Hilda_C_2020_PhD.pdf?sequence=2&isAllowed=y)
- Smith, J. D. et Polaha, J. (2017). Using implementation science to guide the integration of evidence-based family interventions into primary care. *Families, Systems, & Health*, 35(2), 125–135. <https://doi.org/10.1037/fsh0000252>
- Smylie, J. K. (2011). Knowledge translation and indigenous communities : A decolonizing perspective. Dans E. M. Banister, B. J. Leadbeater et E. A. Marshall (dir.), *Knowledge Translation in Context* (p. 181-200). University of Toronto Press.
- Stewart, R. (2019). Using evidence in international development. Dans A. Boaz, H. Davies, A. Fraser et S. Nutley (dir.), *What works now? Evidence-informed policy and practice* (p.171-189). Policy Press.
- Straus, S. E., Tetroe, J. M. et Graham, I. D. (2011). Knowledge translation is the use of knowledge in health care decision making. *Journal of Clinical Epidemiology*, 64(1), 6–10. <https://doi.org/10.1016/j.jclinepi.2009.08.016>
- Straus, S., Tetroe, J. et Graham, I. D. (2013). *Knowledge Translation in Health Care : Moving from Evidence to Practice*. John Wiley & Sons.
- Sussman, S., Pentz, M. A., Rohrbach, L. A., Skara, S. et Valente, T. W. (2006). Translation in the health professions : converting science into action. *Evaluation & the Health Professions*, 29(1), 7–32. <https://doi.org/10.1177/0163278705284441>
- Walter, I., Nutley, S. et Davies, H. (2005). What works to promote evidence-based practice? A cross-sector review. *Evidence & Policy*, 1(3), 335–364.
- Ward, V. L., House, A. O. et Hamer, S. (2009). Knowledge brokering : Exploring the process of transferring knowledge into action. *BMC Health Services Research*, 9(1), 12. <https://doi.org/10.1186/1472-6963-9-12>



# Conclusion





# Futur de l'évaluation : les points de vue des premiers coordonnateur et présidents du Réseau Francophone de l'Évaluation<sup>1</sup>

*Guy Cauquil*

*Thiery Tsou Fematouo*

*Jean Serge Quesnel*

## Introduction

Dans ce chapitre :

Guy Cauquil traite de l'évaluation en mutation. Il expose que l'évaluation n'est pas qu'un simple rapport de rendement des initiatives publiques. C'est un processus réflexif qui engage de plus en plus les parties prenantes dans des processus de transformation développementale menant vers des retombées de biens communs.

Thiery Tsou Fematouo traite du futur de l'évaluation en Afrique. Les tendances récentes pointent positivement vers une utilisation accrue et pragmatique de la fonction d'évaluation. On doit reconnaître l'énorme influence de la culture et de la politique sur la fonction d'évaluation.

Jean Serge Quesnel brosse un tableau de l'émergence de la professionnalisation de l'évaluation. Grâce aux communications électroniques, la professionnalisation se développe à un rythme exponentiel. Déjà, les fondements de la professionnalisation sont en place.

---

1. Jean Serge Quesnel, de la Société québécoise d'évaluation de programme (SQEP) et de la Société canadienne d'évaluation (SCE), fut le premier coordonnateur du Réseau Francophone de l'Évaluation (RFE) (2008-2013). Guy Cauquil, président d'honneur de la Société Française de l'Évaluation (SFE), fut président fondateur du RFE (2013-2016), et Thiery Tsou Fematouo fut le second président de ce réseau.

## **Un futur incertain : l'évaluation confrontée aux mutations culturelles, politiques et éthiques. L'évaluation devra se transformer si elle ne veut pas disparaître.**

L'évaluation est devenue incontournable, parallèlement à la montée en puissance des politiques de développement soutenues par les grands bailleurs internationaux. Or, ceux-ci ont induit une approche très pragmatique de l'évaluation, en raison notamment de :

- la recherche de preuves de l'efficacité de leurs programmes;
- la nécessaire remontée des données de terrain pour éclairer leurs décisions;
- la lutte contre les détournements de l'argent public.

Du coup, ces analyses évaluatives ont avant tout été réalisées au profit des décideurs et des financeurs. Or, les crises économiques, démocratiques et sanitaires des dernières années mettent en évidence la faible efficacité de ces démarches évaluatives confiées aux seuls experts ou aux détenteurs du pouvoir. À cet effet, pensons aux dimensions complexes du développement durable qui oblige une participation plus active des parties prenantes. Ceci impose de repenser le futur de l'évaluation, et cela pour trois grandes raisons :

### **Une crise de légitimité**

Cette crise est générée par les limites d'une approche trop rationnelle de l'évaluation :

- Les limites du modèle du NPM (*New Public Management*) qui a grandement influencé la définition des critères standards de l'évaluation (théorie du changement, démarche projet, rationalisation des moyens, mesure des impacts attendus, etc.) ont montré la relative inadéquation de l'évaluation.
- L'évaluation de type gestionnaire est confrontée à la complexité des décisions politiques. Ce modèle rationaliste ne permet pas de prendre en compte des contingences et des choix complexes propres aux politiques publiques.
- Enfin, cette rationalité des méthodes de prévision et d'évaluation est aujourd'hui remise en cause par la pandémie causée par la maladie à coronavirus 2019, qui impose d'intégrer de nouveaux critères et priorités.

### **Une diversification inéluctable des pratiques évaluatives**

La prise en compte de la diversité culturelle des acteurs de l'évaluation devient nécessaire et passe par :

- la diversité des langues, qui suscite des interrogations quant au modèle originel anglophone de l'évaluation;
- la culture orale de l'évaluation et de la décision publique dans les pays africains, qui requièrent une participation plus active des acteurs de terrain dans l'évaluation.

Les débats sur la fonction démocratique de l'évaluation dans certains pays occidentaux invitent par ailleurs à réfléchir au rapport entre l'évaluation et la décision publique.

## L'émergence de nouveaux critères sociétaux

Ces nouveaux critères invitent à enrichir l'évaluation :

- La prise en compte du genre, de l'équité et du développement durable par les bailleurs internationaux amorce une évolution radicale de l'évaluation. Pensons à l'évaluation *Blue Marble* de Michael Quinn Patton (Patton, 2020).
- L'évaluation de « l'utilité sociale » de l'action publique qui émerge en Europe conduira elle aussi à interroger l'ensemble des critères classiques d'évaluation (Cauquil *et al.*, 2016).

Bref, le futur de l'évaluation dépendra de sa capacité à évoluer. Du simple rapport de rendement des initiatives publiques qu'était l'évaluation, elle évolue vers une démarche qui contribue à une transformation développementale optimisant les retombées de biens communs.

## Le futur de l'évaluation en Afrique : entre inculturation et institutionnalisation

### Une croissance économique remarquable dans un contexte de fragilité

L'Afrique a d'immenses atouts, mais aussi des difficultés chroniques. Avec une population d'environ 1,2 milliard d'individus, un taux de croissance estimé à 3,6 % en 2019 et des ressources naturelles considérables, le continent dispose de solides atouts pour envisager une marche sereine vers son développement (Banque africaine de développement, 2021). Cependant, ces avancées restent fragiles en bien des points. De nombreux conflits sociopolitiques minent le continent. Les disparités entre les couches sociales, notamment entre les sexes, restent tenaces et privent le continent d'une partie de son potentiel de croissance et d'innovation. Environ 416 millions d'Africains vivent toujours dans l'extrême pauvreté. Bien que l'Afrique semble faiblement affectée en apparence par la pandémie de COVID-19, cette dernière risque d'entraîner une récession importante, avec des conséquences incalculables sur la qualité de vie (Banque mondiale, 2021). Les questions de gouvernance restent au cœur des préoccupations majeures, malgré quelques avancées enregistrées ces dernières années. Cette fragilité coûte annuellement un demi-point de croissance au continent (Banque mondiale, 2021).

### L'évaluation en Afrique : entre exercice de conformité et reconnaissance de la nécessité

Il semble évident que l'évaluation trouve toute sa place et son sens dans le contexte d'une Afrique en pleine mutation; entre autres, pour orienter une prise de décisions éclairée par les leçons tirées des décennies d'efforts de développement. Cependant, entre se conformer à un exercice quasi routinier imposé par les procédures de management des projets et reconnaître son utilité, notamment à travers une conception plus pertinente, plus efficace de la relation mandant/mandat/mandataire, se traduisant par une exploitation effective des résultats d'évaluation, il y a un pont à construire pour une évaluation transformatrice. Certaines questions semblent au cœur de cette action :

- Dans sa présentation actuelle, l'évaluation apparaît comme un concept structuré en Occident et arrivé dans les mallettes des bailleurs de fonds en même temps qu'un certain nombre de mesures économique-financières diversement accueillies et dont les résultats (aux finalités souvent jugées énigmatiques) n'ont pas fini d'alimenter les controverses et les polémiques;

- L'examen des processus de fixation des critères/standards d'évaluation doit être fait pour en faciliter l'acceptation et l'appropriation. Il sera intéressant de voir si la recherche se saisira de ces questions dans les prochaines années.

### **Le positionnement de la dimension redevabilité bénéficiaire doit être clarifié versus la redevabilité partenaire. L'évaluation au contact des cultures africaines.**

Le débat sur l'institutionnalisation de l'évaluation traverse le continent de bout en bout. On voit bien que les approches divergent selon les influences culturelles. Mais il semble y avoir unanimité pour reconnaître qu'il s'agit de développer une véritable culture de l'évaluation depuis la population jusqu'à sa codification dans les cadres institutionnels fondamentaux, et d'assurer l'effectivité de son fonctionnement à travers différents mécanismes. En somme, travailler à faire émerger de façon endogène une demande et une offre d'une évaluation au service d'un développement maîtrisé. Cependant, dans le cadre de ces réflexions, il sera utile de sortir du biais d'une conceptualisation occidentale de l'évaluation comme panacée à diffuser dans les « cultures périphériques », pour reconnaître l'existence dans chaque société d'une capacité à définir et à apprécier ses orientations de développement. En effet, l'évaluation n'est pas perçue de la même façon dans les contextes anglophone, arabophone, lusophone et francophone. Plus encore, ces contextes chargés d'histoire masquent de nombreuses réalités culturelles qui influent certainement de façon endogène sur l'appropriation de l'évaluation en Afrique. L'Association Africaine d'Évaluation (AfrEA) et le RFE offrent déjà quelques plateformes d'échanges interculturels sur cette question. Mais, afin de faciliter son appropriation comme instrument d'une gouvernance durable, la recherche devra travailler à proposer des pistes de réflexion pour des approches évaluatives enrichies des éléments culturels africains au cours des prochaines années. Il sera également nécessaire de *se questionner et de repenser les postulats/hypothèses de diffusion de l'évaluation en Afrique*. En effet, au bout d'une vingtaine d'années d'essais diversement appréciables, il est temps d'évaluer le chemin parcouru et de s'interroger sur les canaux traditionnels utilisés jusqu'ici. Il faut peut-être envisager dès à présent la possibilité d'une approche multisectorielle et pluridisciplinaire de diffusion de l'évaluation.

### **Le futur de la professionnalisation de l'évaluation**

« Professionnaliser une activité signifie en définir le pourtour, clarifier les compétences nécessaires à l'exercice de cette activité, établir des normes de pratiques permettant de distinguer un geste professionnel adéquat d'un autre qui ne l'est pas, énoncer les lignes directrices éthiques qui serviront à prendre des décisions difficiles et possiblement offrir un mécanisme de reconnaissance des professionnels-les. Dans nombre de secteurs de l'activité humaine, la professionnalisation a été le passage obligé pour faire en sorte que les professionnels-les fondent leur pratique sur une connaissance suffisante, pour fournir une assurance aux consommateurs-trices des services et ultimement pour augmenter la qualité des actes professionnels. On pensera aux médecins, aux comptables, aux avocats-desquels on exige typiquement une certification. [...] Le débat sur la professionnalisation de la pratique évaluative est en cours à travers le monde. » (Réseau francophone de l'évaluation, 2019).

On peut classer les adhérents et les contestataires de la professionnalisation de l'évaluation en trois groupes. Il y a les « pro », telles les associations d'évaluation qui octroient un titre professionnel en évaluation, comme la SCE, la Société japonaise d'évaluation (JES) et les universités qui offrent des programmes gradués de formation en évaluation. Il y a les « contra », qui s'objectent à une

spécialisation de l'évaluation croyant qu'il s'agit d'un rétrécissement d'une approche scientifique qui est universelle à toutes les disciplines. Pour eux, il serait imprudent de prendre le chemin de la professionnalisation sans davantage de preuves à l'appui. Les risques cités à la professionnalisation sont : la cristallisation de la pratique, l'exclusion des non-experts, l'expropriation d'un processus évaluatif démocratique, la reconnaissance erronée d'une personne comme qualifiée, la transmission d'un pouvoir indu aux organisations ou aux personnes appelées à définir les paramètres de cette professionnalisation, etc. (Réseau francophone de l'évaluation, 2019). L'Association américaine d'évaluation (AEA) est la plus loquace à cet égard. Enfin, il y a les gens hésitants, qui contribuent par ailleurs à habilitier les cinq éléments nécessaires à la professionnalisation. Selon Stockmann et Meyer (2016), ces éléments sont :

- Un marché pour l'évaluation
  - Une demande pour l'évaluation (spécialement au niveau des politiques publiques nationales);
  - Une offre de services d'évaluation spécifique à la demande.
- La professionnalisation de l'évaluation
  - Une institutionnalisation des échanges et une représentation des intérêts (associations, réseaux);
  - Offre de développement de capacités (programmes universitaires de formation et cours).
- L'utilisation de l'évaluation
  - Dissémination des résultats d'évaluation (nombres d'évaluations et publications).

Il n'y a pas de doute qu'il existe un marché pour l'évaluation. La demande est forte et continue d'augmenter. Nous n'avons qu'à cumuler les investissements faits en évaluation par les gouvernements nationaux, par les organisations régionales et par les institutions internationales.

Le total arrive rapidement à un milliard de dollars par année.

L'offre de services est fournie par une multitude des petites, moyennes et grandes entreprises de professionnels qui répondent à la demande d'évaluation. Les associations et réseaux d'évaluation sont nombreux aux niveaux nationaux, régionaux et mondiaux. L'Organisation internationale pour la coopération en évaluation (OICE), qui est une amalgamation mondiale d'organisations volontaires d'évaluation, a répertorié en date de 2020 une liste de plus de 144 organisations, dont 133 sont au niveau national et 11 au niveau international (OICE, 2020). Ces organisations comptent quelque 40 000 praticiens d'évaluation. Leur nombre augmente chaque année.

L'offre de développement de capacité en évaluation est de plus en plus accessible, particulièrement par l'entremise de la formation en mode virtuel. LaVelle et Donaldson (2010) ont recensé les programmes de formation en évaluation aux États-Unis. En 2008, seulement 3 des 48 programmes étaient d'une envergure suffisante, ayant sept offres de cours ou plus en évaluation. Les autres programmes étaient associés avec d'autres disciplines offrant des cours en évaluation basés sur les méthodologies de la recherche. Or, l'offre de programmes gradués a maintenant doublé. Lors d'un atelier de la conférence annuelle de la Société européenne d'évaluation (SEE), Jacques Toulemonde (2012) a brossé un tableau de la formation offerte en Europe. Quatorze des quarante-sept pays de l'Europe avaient des programmes gradués en évaluation.

Les gouvernements ont des régimes institutionnels différents en ce qui trait au positionnement de l'évaluation. Dans certains pays, les commanditaires procèdent eux-mêmes à évaluer leurs interventions, ce qui favorise la mise place d'un minimum de structure nationale chargée de suivre ces évaluations. Certains pays ont mis en place des structures et des politiques assurant que les évaluations soient crédibles, indépendantes et utiles aux prises de décision. Plusieurs pays ont des associations ou des réseaux d'évaluation réunissant des professionnels impliqués dans des évaluations. Rosenstein (2015) a effectué une cartographie des pays selon leur niveau de maturité quant à une législation nationale en matière d'évaluation. En 2015, elle a examiné 109 pays et elle conclut que :

- 27 pays avaient une politique nationale formelle d'évaluation;
- 32 pays avaient une politique nationale non formalisée d'évaluation, mais une pratique d'évaluation bien ancrée;
- 50 pays avaient une documentation démontrant l'existence de facto d'une politique nationale.

Plusieurs pays de l'Europe de l'Est utilisent l'évaluation pour être en conformité avec les règles de l'Union européenne. Similairement, plusieurs pays africains utilisent l'évaluation pour satisfaire les exigences des organisations du système des Nations Unies et des bailleurs de fonds. Sans doute, les progrès dans l'utilisation de l'évaluation sont le résultat d'influences externes. Il est donc évident que les gouvernements et les associations sont en voie de mettre en place les fondements de la professionnalisation de l'évaluation.

En ce qui a trait à l'évolution de l'évaluation à l'échelle mondiale, on peut affirmer que le progrès est lent mais certain. Les politiques nationales d'évaluation qui existent sont flexibles, permettant une adaptation face aux environnements de gouvernance qui changent et à une institutionnalisation graduelle. Elles sont un modèle pour les autres pays qui œuvrent à mettre en place un régime national d'évaluation. Les tendances mondiales pour le futur de l'évaluation sont positives. Le défi est celui de renforcer l'intégration de l'évaluation comme une pratique qui allie les démarches scientifiques et les besoins d'instruments pragmatiques qui répondent aux besoins du politique et du système démocratique. Mondialement, il y a émergence d'une communauté de praticiens d'évaluation qui reconnaît la pluralité des approches à l'évaluation tout en poursuivant la quête de rigueur et la crédibilité. Enfin, l'évidence sera l'utilisation de l'évaluation aux fins de la bonne gouvernance.

## Références bibliographiques

- Banque africaine de développement. (2021). *Perspectives économiques en Afrique de l'Ouest*. <https://www.afdb.org/fr/documents/perspectives-economiques-en-afrique-de-louest-2021>
- Banque mondiale. (2021). *Afrique – vue d'ensemble*. <https://www.banquemonde.org/fr/region/afr/overview#1>
- Cauquil, G., Delalande-Dauzié, C., Guérin, M. et Ravoux, F. (2016). *L'utilité sociale, critère de méta-évaluation dynamique* (Cahiers de la SFE n° 10). Société française de l'évaluation.
- LaVelle, J. M. et Donaldson, S. L. (2010). University-Based Evaluation Training Programs in the United States 1980-2008 : An Empirical Examination. *American Journal of Evaluation*, 31(1), 9-23. <https://doi.org/10.1177/1098214009356022>
- OICE. (2020). *VOPE Directory*. <https://ioce.net/vope-directory>
- Patton, M.Q. (2020). *Blue Marble Evaluation – Premises and Principles*. The Guilford Press.
- Réseau francophone de l'évaluation. (2019). Forum F – Les défis de la professionnalisation dans l'espace francophone [Résumé ]. Forum International Francophone de l'Évaluation (FIFE) 2019. [https://www.rfevaluation.org/FIFE/FIFE2019/Foras#h.p\\_cxrOJkfkA\\_9-](https://www.rfevaluation.org/FIFE/FIFE2019/Foras#h.p_cxrOJkfkA_9-)
- Rosenstein, B. (2015). *Mapping of National Evaluation Policies* (2<sup>e</sup> édition). EvalPartners. DOI :10.13140/RG.2.1.2008.0809
- Stockmann, R. et Meyer, W. (dir.). (2016). *The Future of Evaluation – Global Trends, New Challenges, Shared Perspective*. Palgrave Macmillan.
- Toulemonde, J. (2012). *Building Evaluation Capacity Through University Programmes : Where are the evaluators of the future?* 10<sup>th</sup> EES Biennial Conference [Symposium]. Helsinki.





# Conclusion

« Aucun de nous ne sait ce que nous savons tous, ensemble. »

(Anonyme, s. d.)

Il n'est pas aisé de conclure un ouvrage comme celui-ci, car plus que l'exercice de rédaction d'un manuel, ce fut une découverte, un voyage à travers la diversité des disciplines, des organisations, des horizons et des cultures d'où provient chacune et chacun, tout en ayant en commun leur passion pour cette *transdiscipline* qu'est l'évaluation de programmes (Scriven, 2003). De cette aventure, les auteurs et autrices ont rapporté dans leurs bagages une mosaïque de science, d'expérience, de conscience et d'éthique que nous résumons dans les paragraphes suivants.

## L'évaluation est un outil de gestion, d'amélioration continue et d'apprentissage

Au sein des organisations, l'évaluation et ses compères, la planification et le suivi, constituent un fondement pour la gestion axée sur les résultats, héritée du nouveau management public (Mayne, 2007). Au-delà des outils que nous partageons, il s'agit du développement d'une réelle culture évaluative ancrée dans les processus organisationnels dont il est question. À travers le renforcement des capacités tant individuelles qu'organisationnelles, voire nationales ou internationales, la pensée évaluative s'invite dans les analyses situationnelles et favorise une meilleure compréhension des enjeux et des possibilités concourant à l'atteinte d'objectifs de développement durable. Les structures accompagnatrices, les sociétés nationales et internationales d'évaluation, les experts et expertes, consultantes et consultants et le monde académique manient la boîte à outils de l'évaluation pour collecter des données scientifiquement valides, socialement acceptables et contextuellement adaptées aux réalités locales pour conduire une démarche visant à améliorer les interventions et à soutenir la prise de décision. Une telle démarche ne saurait faire l'économie de l'inclusion des savoirs tacites et implicites, de processus participatifs pratiques et transformateurs, ni de la recherche d'un sens commun et d'une valeur publique. Pour paraphraser Mintzberg lorsqu'il évoque *la gestion dans l'action*, nous pouvons certainement dire que pour obtenir une « bonne » évaluation, vous « rassemblez une bonne quantité d'artisanat, ajoutez-y la juste touche d'art et saupoudrez le tout de la science, vous obtiendrez alors un travail qui est avant tout une pratique qui se maîtrise par l'expérience et qui est ancrée dans le contexte » (Mintzberg, 2010, p.20).

## L'évaluation est une discipline sociale

« Discipline de comportement et de pensée sont inséparables dans la construction de l'être et dans sa socialisation » (Guaino, 2015).

Si nous partons de la prémisse que l'évaluation s'intéresse aux interactions qui structurent les comportements humains, tant au sein des organisations que de la société elle-même; et si nous postulons que ces derniers sont également en interaction avec des interventions visant à résoudre une situation sociale problématique, alors nous pouvons considérer que l'évaluation est une discipline sociale,

c'est-à-dire une branche de savoir développée par une communauté de spécialistes adhérant aux mêmes pratiques. Pour autant, il serait ambitieux de dire que l'évaluation est une science à part entière, c'est-à-dire un ensemble cohérent de connaissances relatives à certaines catégories de faits, d'objets ou de phénomènes obéissant à des lois et/ou vérifiés par les méthodes expérimentales (Patton, 2018a). Sans vouloir sombrer dans la philosophie des sciences, l'évaluation comme discipline garde encore cette latitude d'emprunter aux sciences sociales et naturelles une multitude de cadres théoriques, méthodes et outils de collecte et d'analyse de données pour apprécier la valeur d'une intervention dans ses dimensions sociale, économique, culturelle, scientifique, environnementale, qu'elles soient matérielles ou symboliques (Revillard *et al.*, 2021). En embrassant *le contexte* comme variable d'interaction, l'évaluation ne peut prétendre fournir des réponses qui fonctionnent systématiquement et de manière universelle aux problèmes complexes auxquels sont confrontées les sociétés dans leur diversité. En revanche, elle permet de porter un jugement suffisamment crédible sur la valeur d'une intervention en tentant d'offrir une meilleure compréhension de l'influence des structures, des acteurs et de l'environnement sur celle-ci; que ce soit à travers l'examen de sa pertinence, de sa cohérence, de son efficacité, de son efficience, de ses impacts et de la durabilité de ses effets sur les individus, les organisations, les communautés, voire sur la société dans son ensemble. D'ailleurs, la plupart des interventions ciblées par une démarche évaluative visent à répondre à des enjeux sociaux traditionnels et émergents, en vue d'améliorer les conditions de vie, voire à servir de levier pour la transformation sociale au sein de sociétés plus équitables, prospères et inclusives.

## L'évaluation est une pratique professionnelle

Être évaluateur ou évaluatrice consiste à pratiquer une profession dans laquelle l'on mobilise un ensemble de savoirs, de savoir-faire et de savoir-être. Ainsi, la crédibilité de l'évaluatrice ou de l'évaluateur *professionnel* est tributaire à la fois de sa formation initiale (rarement effectuée en évaluation) et des compétences acquises et actualisées tout au long de son parcours professionnel, mais dépend aussi de sa capacité à faire preuve d'ouverture à d'autres champs disciplinaires et de pratiques. Or, comme cela a été commenté dans des chapitres de l'ouvrage<sup>1</sup>, il n'existe pas d'ordre professionnel régissant le domaine. À défaut d'avoir un cadre déontologique, la communauté a opté pour la crédibilité professionnelle. Néanmoins, face à l'émergence des évaluatrices et évaluateurs ainsi qu'à la demande exponentielle des évaluations, notamment en contexte international, les sociétés nationales d'évaluation dans l'espace francophone se sont investies considérablement dans le développement de normes, de valeurs et de standards éthiques ainsi que dans la promotion d'une pratique réflexive constante. Cela permet à chacun et chacune d'améliorer son expertise tout en favorisant des échanges avec des pairs plus chevronnés et en s'exposant à l'appréciation des mandataires, le cas échéant. Le processus de professionnalisation ne s'est pas encore généralisé au niveau international, mais le mouvement est en marche, notamment au Canada où la Société canadienne d'évaluation a instauré, dès 2015, le titre d'évaluatrice ou d'évaluateur qualifié. Le défi demeure de standardiser une profession perçue parmi les plus démocratiques et inclusives, puisqu'elle accueille dans ses rangs, à l'image de cet ouvrage, des individus aux parcours et aux expériences multiples.

---

1. Notamment, « Principes, normes, règles et standards, éthique et déontologie de l'évaluation » de Quesnel et Bencheikh, ainsi que « Le futur de l'évaluation : le point de vue des trois premiers coordonnateurs du Réseau Francophone de l'Évaluation » de Quesnel, Cauquil et Fematouo.

## L'évaluation est un objet politique

Dans sa thèse intitulée *L'évaluation, un objet politique : le cas d'étude de l'aide au développement*, Laporte (2015) défend l'idée selon laquelle « la mise en pratique de l'évaluation est déterminée par l'environnement politique dans lequel elle intervient, et où elle influence à son tour les politiques des donateurs et des bénéficiaires de l'aide au développement ». Ainsi, les enjeux de l'évaluation sont le reflet des débats de société autour de la santé et de l'éducation en passant par les changements climatiques, le développement durable, les droits de la personne, les questions de genre et d'autonomisation des peuples. En contexte international, ils prennent d'autant plus d'ampleur lorsque l'évaluation porte un jugement sur la valeur de l'efficacité de l'aide ou encore lorsqu'elle contribue à déterminer le niveau de bonne gouvernance des États, et ce, selon les perspectives d'acteurs multiples et entretenant des rapports de force souvent déséquilibrés. Mais l'évaluation en tant qu'objet politique, c'est aussi l'implication des évaluateurs et évaluatrices dans la mise en exergue de ces multiples perspectives, en créant des espaces de dialogue ou de délibération, autant entre des théories qui proviennent principalement du Nord et la réalité des pays du Sud qu'entre différents groupes et décideurs au sein d'une communauté. Lorsque les résultats des évaluations sont diffusés, ils peuvent potentiellement contribuer au débat démocratique en impliquant la société civile et les médias, ce qui peut redonner toute sa place à la participation citoyenne et permettre un certain rééquilibrage des pouvoirs. Peu importe nos positions sur ces sujets, chaque fois qu'une mobilisation sociale ou professionnelle, comme l'Association africaine d'évaluation (AfrEA), s'investit dans des discussions sur la légitimité des modèles d'évaluation en Afrique ou que des auteurs invitent à « décoloniser » la pratique évaluative (Chilisa *et al.* 2017), le débat a le mérite d'être soulevé.

## L'évaluation est une aventure humaine : importance du « facteur personnel »

Les paragraphes précédents démontrent jusqu'à quel point le champ de l'évaluation est dynamique et se réinvente constamment. Ce dynamisme, nous le devons aux femmes et aux hommes qui s'investissent dans ce domaine, qui partagent leurs connaissances, leurs défis et leurs expériences, tout en incitant la jeune génération à prendre sa place sur la scène internationale. Nous aspirons au développement des pratiques par la formation, le mentorat, l'innovation et la créativité, et surtout par les échanges d'expériences en évaluation pour des politiques publiques efficaces, équitables, résilientes et rassembleuses. Ce « facteur personnel », au sens de Patton (2018b), tant prôné dans la communauté des évaluateurs et évaluatrices, est fondé sur notre posture professionnelle. Tout en respectant les normes et standards d'évaluation, nous devons nous interroger sur la posture que nous privilégions en tant que professionnels lorsque nous envisageons de conduire une démarche évaluative.

Ce positionnement est un choix unique, personnel. Il y a autant de postures que d'identités chez les évaluatrices et évaluateurs. Notre posture se construit à partir de **nos valeurs** : le concept le plus central dans notre profession. Celles-ci sont le résultat de notre parcours personnel et professionnel. Elles influencent notre façon de voir le monde, et par conséquent, ce que nous valorisons dans la démarche évaluative, aussi systématique soit-elle. Conscients de ces valeurs, nous devons également reconnaître que celles-ci sont porteuses de biais. Dans le monde réel, nous sommes, nous aussi, en tant qu'évaluateurs et évaluatrices, soumis aux mêmes erreurs, biais méthodologiques, préjugés et pressions que tout autre être humain. Néanmoins, grâce à l'éthique qui guide notre travail, nous avons l'avantage d'en être conscients et surtout de faire le nécessaire pour les corriger. La qualité d'un

bon évaluateur ou évaluatrice ne se mesure pas uniquement à l'aune de sa capacité de planification, des cadres théoriques mobilisés ou des méthodes déployées; elle dépend aussi de sa *sagesse pratique* (House, 2015). Cette notion renvoie à l'idée de « faire la bonne chose en tenant compte des circonstances particulières associées à la réalisation d'une tâche professionnelle » (traduction libre de House, 2015, p. 88). En quelque sorte, faire preuve de sagesse pratique « ne consiste pas à établir des règles et à les suivre [mais] de savoir si, quand et comment [les] appliquer » (traduction libre de House, 2015, p. 90). Elle s'acquiert principalement sur le tas à travers nos interactions humaines et nos expériences tant positives que négatives. La force de la sagesse pratique réside dans une combinaison vertueuse de nos trajectoires personnelles, culturelles et professionnelles. Finalement, notre identité professionnelle se construit à travers notre capacité à trouver le juste équilibre entre nos valeurs et les faits observés; entre nos biais et l'application d'une rigueur méthodologique pour les atténuer.

Nous espérons que cet ouvrage aura réussi à contribuer au champ de l'évaluation dans l'espace francophone tout en faisant émerger autant de réflexions, de créativité que le courage nécessaire pour maintenir notre discipline aussi vivante que résiliente.

## Références bibliographiques

- Chilisa, B., Major, T. E., Gaotlhobogwe, M. et Mokgolodi, H. (2017). Decolonizing and indigenizing evaluation practice in Africa: Toward African relational evaluation approaches. *The Canadian Journal of Program Evaluation*, 30(3), 312–328. <https://evaluationcanada.ca/system/files/cjpe-entries/30-3-313.pdf>
- Guaino, H. (2015, 23 janvier). Discipline de comportement et de pensée sont inséparables dans la construction de l'être et dans sa socialisation. *Le Figaro*. <http://evene.lefigaro.fr/citation/discipline-comportement-pensee-inseparables-construction-etre-s-3333933.php>
- House, E.R. (2015). *Evaluating: Values, biases, and practical wisdom*. Information Age Publishing.
- Laporte, C. (2015). *L'évaluation, un objet politique : le cas d'étude de l'aide au développement* [thèse de doctorat, Institut d'études politiques de Paris]. <https://tel.archives-ouvertes.fr/tel-01164807/document>
- Mayne, J. (2007). Challenges and Lessons in Implementing Results-Based Management. *Evaluation*, 13(1), 87-109. <https://doi.org/10.1177/1356389007073683>
- Mintzberg, H. (2010). *Gérer (tout simplement)*. Transcontinental.
- Patton, M. Q. (2018a). Evaluation Science. *American Journal of Evaluation* 39(2), 183–200. <https://doi.org/10.1177/1098214018763121>
- Patton, M. Q. (2018b). *Facilitating Evaluation: Principles in Practice*. Sage Research Methods. <https://dx.doi.org/10.4135/9781506347592.n4>
- Revillard, A., Delahais, T., Devaux-Spatarakis, A. et Ridde, V. (2021). Introduction : l'évaluation est-elle une science ? Dans T. Delahais, A. Devaux-Spatarakis, Revillard, A. et Ridde, V. (dir.), *Évaluation : fondements, controverses, perspectives*. Éditions des sciences du bien commun.
- Scriven, M. (2003). Evaluation in the new millennium: The transdisciplinary vision. Dans S. I. Donaldson et M. Scriven (dir.), *Evaluating social programs and problems: Visions for the millennium*. Routledge.

# Notices biographiques des auteurs et autrices

**Aref Ben ABDALLAH**, B. Ing, MGP, PMP®, GBLSS, conseiller aux affaires internationales, Direction des affaires internationales, École nationale d'administration publique (ENAP).

M. Abdallah est expert senior avec plus de 30 années d'expérience professionnelle en gestion axée sur les résultats, budget programme, en évaluation de programmes et en planification stratégique. Il a participé, en tant que directeur, chef d'équipe, consultant, formateur et évaluateur à la bonne conduite de nombreux mandats d'envergure. Il a dirigé et animé (en français, en anglais, en arabe) plusieurs panels et formations en évaluation, notamment TAQYEEM et également dans le cadre du Programme international de formation en évaluation du développement (PIFED). Il est un membre actif des discussions sur le développement des capacités en évaluation de l'Initiative globale en évaluation (GEI) et de plusieurs associations d'évaluation. Il détient une maîtrise en gestion de projet de l'Université du Québec à Montréal (UQAM) et en gestion des technologies de l'information de l'Université Laval, certifié en Project Management Professional (PMP) et ceinture verte Lean Six Sigma (CSSC).

**Marie-Hélène ADRIEN**, Ph. D., consultante senior en évaluation, Le Groupe de gestion Universalia, ([www.universalia.com](http://www.universalia.com)).

M<sup>me</sup> Adrien a plus de 30 ans d'expérience en évaluation des organisations et des partenariats auprès d'une clientèle incluant les agences des Nations Unies, la Banque mondiale et le secteur privé philanthropique. Sa carrière l'a amenée à évaluer des projets, des programmes et des politiques publiques dans plus de 50 pays. Elle a enseigné au Programme international de formation en évaluation du développement (PIFED) de la Banque mondiale et à l'Institut d'étude du développement international (IEDI) de l'Université McGill. Ancienne présidente de l'International Development Evaluation Association (IDEAS), elle a aussi siégé sur les conseils d'administration de la Société québécoise d'évaluation de programme (SQEP) et de la Société canadienne d'évaluation (SCE). Auteure de plusieurs ouvrages, dont *L'Amélioration de la performance organisationnelle*, elle s'implique dans la préparation de la relève en évaluation sous forme de mentorat auprès des jeunes évaluateurs émergents.

**Eva ANSTETT**, M.A.P., experte en évaluation.

M<sup>me</sup> Anstett est titulaire d'une maîtrise en administration publique de l'École nationale d'administration publique (ENAP) de Québec, avec une spécialisation en évaluation de programme. Après avoir occupé plusieurs postes liés à la recherche et à l'enseignement de l'évaluation, dont la coordination pédagogique du Programme international de formation en évaluation du développement (PIFED) entre 2015 et 2019, elle exerce actuellement en tant qu'experte indépendante dans le cadre de projets de développement internationaux ou directement auprès d'administrations engagées dans la gestion axée sur les résultats (GAR). Chargée de cours et enseignante en perfectionnement, elle a formé de nombreux administrateurs et acteurs publics sur des thèmes liés à la GAR au Canada et dans les pays de la Francophonie. En parallèle, elle a également mené la réalisation de nombreuses missions d'évaluations et accompagné diverses organisations publiques dans leurs travaux d'évaluation.

**Oumoul Khayri BA TALL**, MBA, experte-comptable et spécialiste en évaluation et en entrepreneuriat, directrice exécutive du cabinet OKT-Consult.

M<sup>me</sup> Ba Tall a fondé le cabinet OKT-Consult en 1997 à Nouakchott (Mauritanie). Elle compte 30 ans d'expérience professionnelle, principalement axée sur les services de gestion/management, de contrôle/*oversight*, et sur la formation professionnelle. Elle a participé au lancement du Réseau Francophone de l'Évaluation (RFE) et intervient dans le Programme international de formation en évaluation du développement (PIFED). Ancienne présidente de l'African Evaluation Association (AfrEA) de 2007 à 2009, et de l'International Organisation for Cooperation in Evaluation (IOCE) de 2008 à 2010, elle siège au conseil d'administration de la International Initiative for Impact Evaluation (3ie) depuis 2012, et du BACoMaB depuis 2019 où elle préside la commission d'audit. Elle est engagée dans de nombreuses initiatives pour renforcer les systèmes d'évaluation, et dans la pratique de l'évaluation du développement en Afrique et dans le monde.

**Ahmed BENCHEIKH**, Ph. D., directeur général du cabinet-conseil Evalua Development et Evalua Institute (Casablanca, Maroc) et professeur associé à l'Université Mohammed VI Polytechnique (Rabat, Maroc).

Spécialiste international en conception, en mise en œuvre et en évaluation des politiques de développement, M. Bencheikh a réalisé plusieurs mandats d'évaluation sur les questions environnementales (biodiversité, changement climatique, etc.), économiques et sociales (économie circulaire, employabilité des jeunes, préscolaire, etc.). Il intervient dans le Programme international de formation en évaluation du développement (PIFED). Il est président-fondateur de l'Association Marocaine de l'Évaluation (AME) (2008), membre fondateur du réseau The Miggle East and North Africa Evaluation Network (EvalMENA) (2009) et du Réseau Francophone de l'Évaluation (RFE) (2009). Il est récipiendaire du Prix Hassan II pour l'environnement (2001), du Prix de l'Association Africaine de l'Évaluation (AfrEA) (2012), de la distinction du réseau EvalMENA (2014) et du Prix du RFE (2022).

**Faya François BOUROUNO**, économiste, consultant international, chef de mission chez Enabel.

M. Bourouno est détenteur d'un master en analyse et évaluation économique des projets. Il a participé au Programme international de formation en évaluation du développement (PIFED) et a obtenu une certification en évaluation d'impact des politiques publiques à l'École nationale de la statistique et de l'analyse de l'information (ENSAI) en France. Il travaille depuis 2012 à la Banque mondiale et a participé à l'évaluation d'impact de plusieurs programmes. Il travaille actuellement sur l'analyse des indicateurs de l'indice du capital humain (ICH). Il cumule plusieurs expériences en évaluation des projets dans plusieurs domaines (éducation, consolidation de la paix, emploi et insertion des jeunes, protection sociale, migration) en Guinée, en République démocratique du Congo, au Burundi et en République centrafricaine. Il appuie actuellement le gouvernement burundais dans l'élaboration du Plan d'action pour l'Emploi. Il a évalué des projets pour le compte de plusieurs organisations du système des Nations Unies et autres organisations internationales. Il dirige le bureau d'études YEMEC et utilise couramment des solutions CAVI/CAPI/BI/Cloud dans l'évaluation des projets.

**Isabelle BRUNET**, M.A.P., chargée d'enseignement, École nationale d'administration publique (ENAP).

M<sup>me</sup> Brunet a développé son expertise lors de mandats d'évaluation locale, régionale, nationale et internationale en environnement, en économie, en éducation et en santé. Elle a complété une formation Lean Six Sigma aux HEC en 2013 et un microprogramme de 2<sup>e</sup> cycle en Gestion de la performance et de l'amélioration continue à l'Université Laval en 2014. Elle est chargée



d'enseignement à l'ENAP depuis 14 ans. Son enseignement, axé sur la pratique, et son rôle de conseillère académique centré sur les besoins du milieu, traduisent son intérêt pour l'application concrète des principes d'évaluation. En 2015, elle a joint l'équipe Performance, analyse et évaluation du Centre intégré universitaire de santé et de services sociaux (CIUSSS) Centre-Sud-de-l'Île-de-Montréal, où elle s'est investie dans le développement d'une culture de l'évaluation. Depuis 2021, c'est en tant que conseillère au développement de la pratique professionnelle qu'elle soutient la mise aux normes et les mécanismes de suivi qui s'y rattachent.

**Diana CARTIER**, cheffe de mission à l'Organisation Internationale pour les Migrations (OIM).

M<sup>me</sup> Cartier est détentrice d'un master en sciences politiques, mention études internationales, Institut d'études politiques (Sciences Po) Paris, France. Elle est également bachelor en sociologie et anthropologie de Macalester College, St-Paul, États-Unis. Elle est experte en évaluation dans le domaine humanitaire et développement. Elle est actuellement cheffe de mission à l'OIM, où elle travaillait auparavant comme *Oversight Officer* en suivi et évaluation.

**Guy CAUQUIL**, Ph. D., premier président du Réseau Francophone de l'évaluation (RFE), après avoir été pendant quatre ans président de la Société Française de l'Évaluation (SFE). Titulaire d'un doctorat en sciences de l'éducation et d'un master en psychologie sociale, il est devenu au fil de sa carrière professionnelle, d'une part un spécialiste de l'évaluation des politiques publiques et, d'autre part, un spécialiste des processus de management et d'organisation. Il a dirigé pendant 20 ans le cabinet d'ingénierie sociale CIRESE, spécialisé dans l'évaluation du traitement des problèmes de société, avant de devenir professeur associé en management public à l'Institut de Sciences Politiques de l'Université de Bordeaux. Il dédie aujourd'hui son temps disponible au développement de l'évaluation dans l'ensemble des pays francophones.

**Christian DAGENAIS**, Ph. D., professeur titulaire au département de psychologie de l'Université de Montréal.

M. Dagenais se spécialise dans l'étude des processus et des effets de différents mécanismes de transfert de connaissances (TC), discipline qu'il enseigne à des étudiants de maîtrise et de doctorat. Depuis 2009, il est à la tête de l'équipe RENARD, une équipe fortement investie dans la recherche sur le TC dans le domaine social, et dont l'objectif est de développer ce champ de recherche grâce à des partenariats interdisciplinaires et interuniversitaires. Au cours des 20 dernières années, il a dirigé de nombreuses études évaluatives portant sur le transfert des connaissances issues de la recherche en éducation, en santé publique, en santé mondiale et dans plusieurs autres domaines au Québec et à l'étranger. Il est rédacteur en chef de la *Revue TUC*, une revue francophone de recherche sur le transfert de connaissances <https://revue-tuc.ca/index.php/accueil>

**Thomas DELAHAIS**, cofondateur de la société coopérative Quadrant Conseil (Paris, France).

M. Delahais évalue depuis près de 20 ans des interventions à tous les niveaux de l'action publique, des projets aux politiques publiques et du local à l'international. Il accompagne des démarches d'évaluation, de capitalisation et de (re)conception des interventions, et forme des agents publics et des porteuses et porteurs de projet à l'évaluation et à la conception. Ses travaux portent sur plusieurs sujets, notamment : le pluralisme des approches pour évaluer les interventions complexes et en particulier l'analyse de contribution, qu'il a contribué à opérationnaliser; l'évaluation des initiatives de

transformation et de transition socioécologique et leur accompagnement; les façons de soutenir la posture évaluative parmi les non-évaluatrices et non-évaluateurs; et enfin la sociologie de l'évaluation. Il est membre du bureau éditorial du *Evaluation Journal*.

**Maxime DESFOSSÉS**, B.A., assistante de recherche à la Direction des affaires internationales, à l'École nationale d'administration publique (ENAP).

M<sup>me</sup> Desfossés est titulaire d'un baccalauréat en sociologie. Elle est actuellement étudiante à la maîtrise en sociologie, avec mémoire. Son mémoire porte sur les effets de la COVID-19 sur les travailleurs et travailleuses du sexe au Québec et en France. Elle travaille en collaboration avec le Collectif pour un Québec sans pauvreté, pour l'Université Laval ainsi que pour Services Québec. Elle a reçu la mention d'honneur du doyen pour les années 2019-2020 ainsi que la bourse d'admission à la maîtrise Jean-Charles Falardeau à la suite de l'examen de son dossier par un jury.

**Kris Essosola EALE**, directeur général de ConsultHuntLink Africa.

Avant de fonder ConsultHuntLink Africa, M. Eale a travaillé pendant plus de six ans pour Khulisa Management Services, où il a occupé le poste de spécialiste du Suivi, de l'Évaluation, de la Recherche et de l'Apprentissage (SERA) en santé publique, mais il a aussi dirigé les études de SERA en énergie et développement social. Il est titulaire d'un graduat en médecine physique et réadaptation de l'Université de Kinshasa en République démocratique du Congo (RDC) et d'une maîtrise en santé publique de l'Université d'Afrique du Sud. Il a travaillé pendant plus d'une décennie dans le secteur du suivi, de l'évaluation et de la recherche où il a participé et dirigé des évaluations, des recherches SERA dans plus de 17 pays d'Afrique et à Haïti, dans les Caraïbes. Il avait travaillé auparavant comme agent de terrain pour le programme de vaccination et de lutte contre les maladies transmissibles de l'enfant en RDC.

**Thierry Tsou FEMATOOU**, ingénieur/M. Sc., chargé de programme suivi et évaluation UNWOMEN.

M. Fematouo est ingénieur agro-socio-économiste de l'Université de Dschang; il est titulaire d'un master en gouvernance et management des politiques publiques (option commande publique) co-organisé par Sciences Po/Paris – Université de Turin et ITC-ILO – et d'une certification en évaluation des projets (Carleton University/IPDET). Il a conduit des missions d'expertise sur l'analyse des actions de développement et a élaboré des stratégies pour de multiples partenaires bilatéraux ou multilatéraux. Il est enseignant associé dans plusieurs cursus professionnels et universitaires sur la gestion axée sur les résultats (GAR). Il est l'ancien président du Réseau Francophone de l'Évaluation (RFE) et secrétaire général de la Cameroon Development Evaluation Association (CaDEA). Ses principaux champs de compétences sont l'appui aux processus d'institutionnalisation de l'évaluation, l'analyse des politiques publiques et des stratégies de développement dans le Sud et la gestion des connaissances et des partenariats. Il est lauréat du Prix pour la contribution au développement de l'évaluation dans le monde francophone.

**Alexandre FORTIN**, M.A.P., conseiller en évaluation de programme, Secrétariat du Conseil du trésor (SCT).

Bachelier en économie et politique de l'Université de Montréal et diplômé d'une maîtrise en administration publique de l'École nationale d'administration publique (ENAP), M. Fortin a intégré la fonction publique québécoise en 2021 en tant que conseiller en évaluation de programme. Il a rejoint le SCT au sein de la direction responsable de l'application de la Directive concernant l'évaluation de programme dans les ministères et les organismes. Il a aussi occupé pendant plusieurs années un emploi

d'analyste médiatique au sein de Référence Média, entreprise spécialisée en veille médiatique. Enfin, il s'est impliqué bénévolement dans la vie démocratique, notamment en occupant des postes au sein d'instances d'un parti politique. Il a été, en 2021, le 2<sup>e</sup> lauréat (cycle maîtrise) du Concours de persévérance scolaire de l'Association étudiante de l'École nationale d'administration publique (AEENAP).

**Christophe FRANZETTI**, M.A., chargé d'évaluation à l'Organisation Internationale pour les Migrations (OIM).

M. Franzetti détient un master en sciences politiques avec mention internationale de l'Université de Genève et de l'Institut des Hautes Études Internationales. Il est un expert en évaluation dans le domaine humanitaire et de développement.

**Mahmoud GHOUIL**, conseiller des services publics et expert-formateur international en planification stratégique, suivi & évaluation (S & É) des programmes de développement selon les normes onusiennes.

M. Ghouil a plus de 20 ans d'expérience avec le gouvernement tunisien (ministère des Finances), et auprès des Nations Unies (ONU), où il a été chargé de la planification et du S & É de la programmation commune du système des Nations Unies en Tunisie. Il est également formateur à l'École nationale d'administration publique (ENAP) (Québec) et consultant auprès de plusieurs organisations internationales. Il est économiste de développement de formation; diplômé de l'Université de Tunis et du Centre d'études et de recherches sur le développement international de l'Université Clermont Auvergne. Il est également diplômé de l'École nationale d'administration (ENA) en Tunisie et en France. Il est titulaire du prix de la Banque mondiale sur l'évaluation des « systèmes d'incitations à l'investissement en Tunisie » et auteur de plusieurs articles dans le domaine de la planification, du S & É et des objectifs de développement durable (ODD). Il est membre actif de certains réseaux internationaux d'évaluation.

**Anna GUEYE**, MGP, Ph. D. (c), ÉQ/CE, évaluatrice accréditée par la Société canadienne de l'évaluation (SCE).

M<sup>me</sup> Gueye est présidente de Services Conseils Gesteval Inc., consortium international de consultants ayant pour mission de contribuer à la gestion et à l'efficacité des pratiques organisationnelles. À titre de coordonnatrice scientifique pour le Programme international de formation en évaluation du développement (PIFED) 2012-2013, M<sup>me</sup> Gueye collabore comme formatrice en Évaluation des politiques publiques, en Gestion axée sur les résultats (GAR) et en Management des risques, depuis 2013 pour la Direction des affaires internationales (DAI) de l'École nationale d'administration publique (ENAP). Elle est également chargée d'enseignement en Gestion du changement, en Management stratégique et en Évaluation de programmes à l'ENAP. Anna Gueye est doctorante en administration scolaire, volet enseignement supérieur à l'Université du Québec à Rimouski et chargée d'enseignement en Gestion du changement.

**N'Diamé GUEYE**, Ph. D., conseiller en formation à l'Université Laval et consultant chez Gesteval Services Conseils Canada.

M. Gueye est un spécialiste des communautés de pratique et de la négociation de sens. Il a développé une expertise reconnue en matière de codéveloppement professionnel et de transfert des connaissances. À titre de chercheur, il a accompagné la mise en œuvre du plan d'action pour l'amélioration du français (PAAF) à l'enseignement primaire et secondaire au ministère de l'Éducation, des Loisirs et du Sport du Québec.

**Carla-Anide GUILLAUME**, chargée de dossier commercial à Affaires mondiales Canada (AMC), au sein de la Division des relations bilatérales avec l'Afrique de l'Ouest et du Centre.

M<sup>me</sup> Guillaume travaille également comme assistante de recherche au Programme international de formation en évaluation du développement (PIFED) et candidate à la maîtrise en administration publique de l'École nationale d'administration publique (ENAP). Elle détient un diplôme de premier cycle en science politique de l'Université Concordia.

**Deo-Gracias HOUNDOLO**, M. Sc., spécialiste senior en évaluation, coordinateur régional du Programme ouest-africain de renforcement des capacités et d'évaluation d'impact, à l'Initiative Internationale pour l'Évaluation d'Impact.

M. Houndolo est membre fondateur de l'Institut Africain de Suivi et Évaluation d'Impact des Politiques, Programmes et Projets de Développement. Il a un diplôme d'ingénieur agronome de l'Université d'Abomey-Calavi, un M. Sc. en développement agricole de l'Université de Copenhague, et a achevé le microprogramme de deuxième cycle en économie appliquée du développement de l'Université Laval. Il est doctorant en économie du développement à l'Institut international d'études sociales de l'Université Erasmus de Rotterdam. Son expertise couvre l'analyse des politiques publiques, l'évaluation d'impact ainsi que la conception et la numérisation des systèmes de suivi et d'évaluation intégrée et dynamique. Il a travaillé pour la Banque mondiale, pour le Fonds international de développement agricole (FIDA), pour la Fondation Mastercard, pour la African School of Economics et pour l'École Nationale d'Administration de l'Université d'Abomey-Calavi.

**Tigawendé Prosper KABORE**, agent de planification, de programmation et de recherche au Centre intégré de santé et de services sociaux (CISSS) de Laval (Canada).

M. Kaboré est titulaire d'une maîtrise en sciences économiques et gestion. Il a suivi une formation de master en gestion des projets et une licence professionnelle en statistiques sociales (Burkina Faso) et a complété la maîtrise avec recherche en administration publique de l'École nationale d'administration publique (ENAP) du Québec. Il possède plus de six ans d'expérience dans les organisations internationales (Save the Children, Danish Refugee Council, Catholic Relief Services, Bureau international des droits des enfants) en tant que spécialiste en suivi, évaluation, redevabilité et apprentissage. Il a été associé à plusieurs mandats de consultations en évaluation et en capitalisations en tant que consultant associé ou spécialiste en gestion des données.

**Bobo B. KABUNGU**, Ph. D., directeur scientifique du Centre de Recherche en Sciences Humaines (CRESH).

M. Kabungu est postdoctorant en développement international. Il enseigne en République démocratique du Congo (RDC), à l'Institut National du Travail Social (INTS), à l'École Nationale d'Administration (ENA), à l'Université de Goma (UNIGOM) et au Tronc commun des banquiers organisé par la Banque Centrale du Congo où il est cadre de direction. Économiste de carrière, formé notamment par le Fonds Monétaire International (FMI), il est détenteur d'un doctorat en administration publique (distinction *Magna Cum Laude*), d'un master 2 en évaluation et management des politiques sociales, de deux diplômes d'études supérieures spécialisées (D.E.S.S.), respectivement en économie de développement et en gestion de projets, ainsi que d'un certificat universitaire en évaluation du développement obtenu à l'École nationale d'administration publique (ENAP) du Québec pour laquelle il a été honoré de collaborer, comme invité, au module sur l'essentiel de l'évaluation en février 2021.

**Kadidiatou KADIO**, Ph. D., M.A.(ps), M. Sc., chargée de recherche à l'Institut de Recherche en Sciences de la Santé (IRSS) du Centre National de la Recherche Scientifique et Technologique (CNRST) du Burkina Faso. Elle est aussi chercheuse associée de l'Équipe Renard, Université de Montréal.

M<sup>me</sup> Kadio est titulaire d'un doctorat en sciences humaines appliquées de l'Université de Montréal. Ses intérêts de recherche portent sur les questions de protection sociale, les politiques et systèmes de santé, les inégalités et l'équité d'accès aux services de santé, le renforcement du pouvoir d'agir, le partage et l'utilisation de connaissances. Elle a également une expérience solide de gestionnaire en développement international et en développement communautaire.

Faisant partie du programme Emerging Voices for Global Health (EV4GH) en 2012, elle a été nommée en 2018 dans la liste des 204 femmes qui parlent et travaillent en français en santé mondiale (Women in Global Health francophone).

**Ousséni KINDA**, directeur du suivi et évaluation (*Country Head of M & E*), Norwegian Refugee Council (NRC), Nigéria.

M. Kinda est économiste et évaluateur, avec une expérience avérée dans la gestion de systèmes de suivi et d'évaluation participatifs. Il a occupé, entre autres, des postes de chercheur et de directeur en suivi et évaluation en assurant un leadership technique. À ce titre, il a travaillé avec des institutions telles que le Fonds international de développement agricole (FIDA), Save the Children, le Fonds des Nations Unies pour l'enfance (UNICEF), le Conseil national de recherches Canadale Norwegian Refugee Council (NRC) ainsi que des organismes gouvernementaux et universitaires. Ses travaux récents se sont focalisés sur l'utilisation de la technologie dans les systèmes de suivi et évaluation, le renforcement de capacités en évaluation axée sur le genre, les enfants et les jeunes et sur les approches d'évaluation autochtones en Afrique. À travers différentes plateformes d'évaluation, il a travaillé sur des approches telles que le mentorat afin de relever les défis auxquels sont confrontés les jeunes africains en évaluation. Il est membre de divers réseaux d'évaluation, auteur et coauteur d'articles et de livres.

**Papa Souleymane KOULIBALY**, responsable de l'appui aux projets (*Project Support Officer*), Organisation Internationale pour les Migrations (OIM), Bureau régional Afrique de l'Ouest et du Centre.

M. Koulibaly est titulaire d'un master professionnel en développement durable et d'un diplôme d'ingénieur des travaux de l'aménagement du territoire et de la gestion urbaine à l'Université Cheikh Anta Diop de Dakar. Il tire profit de plusieurs années d'expérience dans le management, l'évaluation de projets et de programmes. Il a été assistant au Cabinet d'évaluation de politiques publiques IFE (Ingénierie, Formation, Évaluation) avant de rejoindre Innovations Environnement Développement (IED) Afrique au poste de responsable pays de suivi évaluation et gestion des connaissances au Sénégal. Il a été responsable régional de la U.S. Agency for International Development (USAID) en qualité de gestionnaire des connaissances en sécurité alimentaire. Depuis mai 2021, il s'occupe des questions de migration liée à l'environnement et au changement climatique à l'Organisation Internationale pour les Migrations (OIM), Bureau régional Afrique de l'Ouest et du Centre. Il s'active pleinement dans les associations et réseaux d'évaluateurs, comme l'Association Sénégalaise d'Évaluation, l'Association Africaine d'Évaluation et le Réseau Francophone de l'Évaluation (RFE).



**Julie LANE**, Ph. D., professeure spécialisée en transfert des connaissances, Département d'études sur l'adaptation scolaire et sociale de la Faculté d'éducation de l'Université de Sherbrooke.

M<sup>me</sup> Lane est également directrice du Centre RBC d'expertise universitaire en santé mentale. Elle possède un doctorat en éducation, et elle s'est intéressée à la posture d'accompagnement. Elle a œuvré à titre de gestionnaire de projets pendant 15 années dans le réseau de la santé et des services sociaux où elle a développé son expertise en transfert des connaissances. Elle a obtenu des prix et des distinctions au cours de sa carrière. Récemment, en 2021, elle a été lauréate du Mérite estrien pour ses projets en santé mentale jeunesse. De plus, en 2019, elle a reçu le Prix méritas – Hommage en prévention du suicide de l'Association québécoise de prévention du suicide pour sa contribution et son engagement en prévention du suicide.

**Sylvie LEFRANCOIS**, M.A.P., cofondatrice de la firme Sogémap inc.

Depuis 1993, M<sup>me</sup> Lefrançois agit à titre de spécialiste en évaluation de programme auprès de ministères et d'organismes du gouvernement du Québec, du gouvernement du Canada et d'organismes à but non lucratif. Elle est également initiatrice du réseau Sogémap – Afrique, qui vise à offrir du soutien et de l'accompagnement aux organisations gouvernementales de l'Afrique francophone. Elle a réalisé de nombreux mandats portant sur l'évaluation de programme, l'élaboration de cadres (devis) d'évaluation, la formation et l'accompagnement de gestionnaires et de professionnels à la fonction d'évaluation de programme et la réalisation d'enquêtes auprès de diverses clientèles de programmes. Depuis 20 ans, elle est également chargée d'enseignement à l'École nationale d'administration publique (ENAP) et à l'Université du Québec à Montréal (UQAM), où elle enseigne les principes et la méthodologie de l'évaluation de programme, les méthodes quantitatives et l'éthique professionnelle en évaluation.

**Juan Manuel LOPEZ MEJIA**, M.A., M.A.P.

Analyste de politiques au sein de la division des politiques et des négociations multilatérales chez Affaires mondiales Canada (AMC). Il est détenteur d'une maîtrise en sciences sociales en France et d'une maîtrise en administration publique, avec concentration internationale, de l'École nationale d'administration publique (ENAP) de Montréal.

**Richard MARCEAU**, Ph.D., professeur associé à l'École nationale d'administration publique (ENAP).

Codirecteur - responsable de la recherche et du développement de la terminologie en évaluation des politiques. Projet de dictionnaire multilingue de l'évaluation (DME). M. Marceau est titulaire d'un doctorat en sciences politiques de l'Université Laval, d'une maîtrise (M. Sc.) en sciences de l'eau de l'Institut national de la recherche scientifique (INRS) ainsi que d'un baccalauréat (B. Sc.) en physique de l'Université Laval. Depuis 2015, il est codirecteur du DME. Il a fait carrière comme professeur titulaire à l'ENAP. Il a contribué à la création du programme de doctorat de l'ENAP dont il a été le directeur. Il a créé et dirigé le Centre de recherche et d'expertise en évaluation (CREXE) de l'ENAP. Il a réalisé des travaux de recherche qui ont mené à la publication du *Dictionnaire terminologique de l'évaluation* (2014) dont il est coauteur avec Francine Sylvain.

**Esther MC SWEEN-CADIEUX**, psychologue, Ph. D., professeure associée, Département d'études sur l'adaptation scolaire et sociale, Faculté d'éducation, Université de Sherbrooke.

Titulaire d'un doctorat en psychologie, recherche et intervention de l'Université de Montréal en 2019, M<sup>me</sup> Mc Sween-Cadieux s'est intéressée, dans le cadre de sa thèse, à la question de l'utilisation des résultats de recherche dans le domaine de la santé publique au Burkina Faso, en Afrique de l'Ouest.

Ses intérêts de recherche portent sur l'évaluation des processus et des effets des stratégies de transfert de connaissances ainsi que sur la mise en œuvre de mécanismes d'intermédiation tels que le courtage de connaissances. En tant que chargée de projets et consultante, elle est appelée à développer des stratégies et des outils de transfert de connaissances et à offrir divers ateliers de formation dans le domaine.

**Nicolas MONTIBERT**, consultant spécialisé en évaluation et développement de projets.

M. Montibert est géographe et agroenvironnementaliste de formation. Il travaille dans le secteur de la coopération internationale depuis presque 20 ans, principalement en Afrique, en Amérique latine et dans les Caraïbes. Après avoir occupé plusieurs fonctions aux Nations Unies et pour l'organisation non gouvernementale (ONG) Oxfam, il travaille désormais pour de nombreuses ONG et divers bailleurs canadiens et internationaux (Nations Unies, Banque mondiale, Union européenne, Organisation de coopération et de développement économiques (OCDE), etc.). Son travail porte sur diverses thématiques relatives à l'adaptation aux changements climatiques, la sécurité alimentaire, la conservation de la biodiversité et la gestion des ressources naturelles, forestières et côtières, mais également l'eau-hygiène-assainissement (WASH) en situation d'urgence humanitaire et dans le secteur de la santé. Sa méthodologie de travail repose sur la gestion axée sur les résultats selon une approche communautaire qui valorise les savoirs traditionnels locaux, dans une perspective de respect des droits de la personne, d'égalité des genres et de justice climatique.

**Ada OCAMPO**, présidente, IDEAS International.

M<sup>me</sup> Ocampo est une sociologue péruvienne qui a obtenu, avec distinction, une maîtrise en planification et management du développement de l'Université de Wales. Elle a travaillé au sein du système des Nations Unies avec le Fonds international de développement agricole (FIDA), le Fonds des Nations Unies pour la population (FNUAP), le Programme des Nations Unies pour le développement (PNUD), le Fonds des Nations Unies pour l'enfance (UNICEF) et le *United Nations System Staff College* (UNSSC), dans divers pays en Afrique, en Amérique latine, en Amérique du Nord, en Europe et en Asie. Elle a œuvré principalement dans le domaine du réseautage et du développement des capacités en évaluation. Elle a enseigné à Carleton (Ottawa), à la Facultad Latinoamericana de Ciencias Sociales (FLACSO) (Santiago), à la Indian School of Business (Hyderabad) et à la BRAC University (Dhaka). Elle est autrice de plusieurs articles et livres. Elle compte parmi les fondatrices de l'Organisation internationale de coopération en évaluation (OICE) et du ReLAC (Réseau latino-américain d'évaluation). Elle a été une membre clé de EvalPartners, a été coprésidente de EVALSDGs et a appuyé le secrétariat du Forum global des parlementaires pour l'évaluation. Ada est polyglotte, ayant vécu au Costa Rica, aux États-Unis d'Amérique, au Guatemala, en Grande-Bretagne, en Italie et en Thaïlande.

**Miché Ouédraogo**, Ph. D., spécialiste en suivi et évaluation à la Banque mondiale au sein du Groupe d'évaluation indépendante (IEG).

M. Ouédraogo est titulaire d'un doctorat en administration publique option analyse et management des politiques publiques de l'École nationale d'administration publique (ENAP) Québec, d'un master 2 en management des projets de l'Institut africain de management (IAM), d'une maîtrise en économie et sciences de gestion de l'Université de Koudougou (UK) et d'une licence en statistique de l'Institut supérieur des sciences de la population (ISSP) de l'Université de Ouagadougou (UO), au Burkina Faso. Ses travaux de recherche portent essentiellement sur l'évaluation des politiques et



la gestion des connaissances dans le secteur public. Il est auteur d'un livre et coauteur de plusieurs articles scientifiques et de chapitres de livres. de la formation Il intervient également comme chargé de cours à l'ENAP de Québec et au Centre Africain d'Études Supérieures en Gestion (CESAG).

**Jean Serge QUESNEL**, ÉQ/CE, membre titulaire de la Société canadienne d'évaluation (SCE), professeur associé à l'École nationale d'administration publique (ENAP) et professeur invité à l'École du développement international et de la globalisation de l'Université d'Ottawa.

M. Quesnel a été président du Groupe d'experts en évaluation du Comité d'aide au développement (CAD) de l'Organisation de coopération et de développement économiques (OCDE), président fondateur de Groupe de coopération en évaluation (ECG) des institutions financières internationales, président du groupe de rédaction des normes et standards pour l'évaluation au sein du Système des Nations Unies et coordonnateur fondateur du Réseau Francophone de l'Évaluation (RFE). Il est le premier membre officiel de la International Development Evaluation Association (IDEAS). Il est membre titulaire (*Fellow*) et évaluateur qualifié (ÉQ) de la Société canadienne d'évaluation et président de la Société québécoise d'évaluation de programme. L'Organisation internationale de coopération en évaluation (OICE) et *EvalPartners* lui ont décerné une reconnaissance pour sa contribution au développement de la profession d'évaluation à l'échelle mondiale.

**Lynda REY**, Ph. D., professeure en évaluation de programmes à l'École nationale d'administration publique (ENAP) Montréal et professeure associée au Centre Africain d'Études Supérieures en Gestion du Sénégal.

M<sup>me</sup> Rey combine une expertise sur l'utilisation des connaissances scientifiques et l'application de diverses approches évaluatives pour améliorer la pratique et les politiques publiques, notamment en santé publique à l'international. Elle a également occupé le poste de directrice du suivi-évaluation à la fondation canadienne One Drop pour un accès durable à l'eau potable. Diplômée de l'Institut d'études politiques d'Aix-en-Provence et d'un master en politiques de développement et action humanitaire de la Sorbonne en France, elle est titulaire d'un doctorat en santé publique et d'un diplôme de 3<sup>e</sup> cycle en analyse et évaluation des interventions en santé de l'Université de Montréal. Elle détient les titres d'évaluatrice qualifiée de la Société canadienne d'évaluation et d'administratrice agréée du Québec. Elle siège sur le conseil d'administration du Réseau francophone de l'Évaluation (RFE).

**Valéry RIDDE**, Ph. D., directeur de recherche au Centre Population et Développement (CEPED) (<http://www.ceped.org>), une unité de recherche commune à l'Université de Paris et à l'Institut de recherche pour le développement (IRD).

M. Ridde est actuellement basé à l'Institut de la santé et du développement (ISED) de l'Université Cheikh Anta Diop de Dakar (Sénégal). Ses travaux de recherche et d'évaluation portent sur la couverture sanitaire universelle, le financement des services de santé, l'évaluation et le transfert des connaissances.

**Élodie C. ROY**, MGP, conseillère aux affaires internationales, École nationale d'administration publique (ENAP).

Conseillère en gestion axée sur les résultats (GAR), suivi et évaluation au sein d'une firme privée pendant huit ans, M<sup>me</sup> C. Roy a développé et animé plusieurs formations sur mesure visant le développement et le renforcement des compétences pour la planification, la mise en œuvre, le suivi et l'évaluation d'interventions de développement auprès de clientèles adultes à l'échelle internationale, en plus d'avoir participé à des évaluations pour une variété d'organisations nationales, régionales et internationales.

Passionnée par la visualisation des données, elle se distingue par sa créativité et ses compétences pour la préparation de rapports mettant en évidence les réalisations, les bonnes pratiques et les leçons apprises en vue de favoriser leur utilisation.

Ouverte sur le monde, elle a vécu et travaillé en Amérique du Nord, en Amérique latine, en Afrique, en Asie, en Europe et au Moyen-Orient.

**Abdoul Aziz SALL**, M.A.P., formateur associé à l'École nationale d'administration publique (ENAP).

M. Sall est diplômé de l'ENAP, titulaire d'une maîtrise et d'un diplôme d'études supérieures spécialisées (D.E.S.S.) en administration publique/évaluation de programmes. Il a contribué au renforcement de la « fonction évaluation » dans différents pays d'Afrique, dont le Mali, où il a été formateur en évaluation des politiques publiques pour les institutions chargées de l'évaluation des politiques publiques : le Bureau du Vérificateur Général (BVG), la Section des Comptes de la Cour Suprême (SCCS) et le Contrôle Général des Services Publics (CGSP). Il est l'auteur du manuel d'évaluation des politiques publiques de la SCCS du Mali. Il a été formateur en évaluation des politiques publiques pour d'autres institutions supérieures de contrôle (ISC), dont la Cour des Comptes du Sénégal (CCS), pour le compte de l'ENAP. À titre de praticien, il a conduit des évaluations dans plusieurs pays africains.

**Vénétia SAUVAIN**, M. A., conseillère aux affaires internationales, École nationale d'administration publique (ENAP).

Sociologue de formation, doublement titulaire d'un master de l'Institut d'études politiques de Lyon et d'une maîtrise en développement international de l'Université Laval, elle combine des expériences en renforcement des capacités au sein de l'Institut Interaméricain de coopération pour l'agriculture (IICA) et d'institutions publiques péruviennes en partenariat avec la coopération allemande. Au sein du Centre de recherche et d'expertise en évaluation (CREXE), elle a évalué plus d'une douzaine de programmes de l'administration publique et du milieu communautaire québécois dans les domaines de l'éducation, des milieux autochtones, de l'implantation de l'analyse différenciée selon les sexes et de la collaboration interorganisationnelle pour l'aide aux victimes d'actes criminels et la lutte contre les gangs de rue. Elle est actuellement formatrice au sein du Programme international de formation en évaluation du développement (PIFED), qu'elle a contribué à concevoir, et elle agit comme coordonnatrice du présent manuel.

**Nebnoma Kader Arnaud SAWADOGO**, conseiller expert en analyse de la performance organisationnelle, Revenu Québec.

M. Sawadogo est titulaire d'une maîtrise en affaires publiques spécialisée en analyse et évaluation de l'Université Laval, et il cumule plusieurs années de pratique évaluative. Son parcours professionnel l'a notamment conduit à travailler comme agent de recherche au Centre de recherche et d'expertise en évaluation (CREXE) et à titre de consultant pour l'École nationale d'administration publique du Québec (ENAP). Il a réalisé des évaluations de projets de développement en santé et en éducation au Burkina Faso avec l'ONG Oxfam. Il y a par ailleurs offert divers ateliers de renforcement des capacités en suivi-évaluation et en gestion axée sur les résultats au profit d'organisations de la société civile œuvrant en faveur de l'autonomisation économique des femmes et de la participation citoyenne des jeunes. Il est actuellement conseiller expert en analyse de la performance organisationnelle à Revenu Québec où il a également occupé le poste d'évaluateur de programme.

**Mary Pat SELVAGGIO**, directrice de la division de santé de Khulisa Management Services depuis 1998.

Experte en santé publique, en suivi et évaluation et recherche, M<sup>me</sup> Selvaggio dirige les activités de sa division à Khulisa dans toute l'Afrique et ailleurs. Elle se concentre sur les évaluations, les évaluations de performance et les audits de la qualité des données, principalement dans le secteur de la santé. Elle a été responsable d'évaluation sur des projets de suivi et évaluation (S & E) en Afrique. Elle a pris un congé sabbatique de 2004 à 2005 pour mettre en place les bureaux de pays d'Afrique du Sud et du Swaziland pour la fondation Elizabeth Glaser Pediatric AIDS. Avant de rejoindre Khulisa en 1998, elle était responsable de la santé pour la U.S. Agency for International Development (USAID) dans les missions du Zimbabwe, du Mozambique et du Swaziland de 1985 à 1998. Elle est titulaire d'une maîtrise en santé publique de l'Université du Minnesota aux États-Unis.

**Francine SYLVAIN**, D.E.S.S.

Codirectrice - responsable de la recherche et du développement de la méthodologie terminologique. Projet de dictionnaire multilingue de l'évaluation (DME).

M<sup>me</sup> Sylvain est titulaire d'un diplôme d'études supérieures spécialisées en administration publique de l'École nationale d'administration publique (ENAP) et d'un baccalauréat (B.A.) en langue française et rédaction professionnelle de l'Université Laval. Depuis 2015, elle est codirectrice du Projet de dictionnaire multilingue de l'évaluation (DME) et responsable de la recherche et du développement de la méthodologie terminologique. Elle a fait carrière à l'ENAP, où elle a réalisé des travaux de recherche qui ont mené à la publication du *Dictionnaire terminologique de l'évaluation* en 2014 dont elle est coauteure avec Richard Marceau. De plus, elle a agi à titre d'attachée d'administration au Centre de recherche et d'expertise en évaluation (CREXE) où elle a exercé des fonctions de réviseure linguistique et a contribué à des travaux d'écriture dans le champ de l'évaluation.

**Michele TARSILLA**, Ph. D., conseiller régional pour l'évaluation, Bureau régional de l'UNICEF pour l'Afrique de l'Ouest et du Centre (WCARO).

M. Tarsilla est responsable d'assurer la qualité, l'indépendance et l'utilisation de l'évaluation dans les 24 pays de la région. Il possède 20 ans d'expérience dans la conduite d'évaluations sur le terrain ainsi que dans l'organisation des programmes de renforcement des capacités en évaluation dans plus de 70 pays (pas seulement dans l'Afrique subsaharienne, mais aussi en Amérique latine et en Asie). Il a travaillé en six langues (y compris l'espagnol, le portugais et le français) pour plus de 40 agences internationales et de nombreux gouvernements nationaux. Il est titulaire d'un doctorat en évaluation interdisciplinaire à l'Université de Western Michigan (É.-U.) et d'un M.S. en développement international en tant que boursier Fulbright à l'École de service extérieur de l'Université de Georgetown (É.-U.). Il est aussi rédacteur associé de la *Revue Africaine de l'Évaluation*.

**Serge Eric YAKEU DJIAM**, ingénieur, M. Sc., ÉQ/CE ; expert international, vice-président de l'Association Internationale pour le Développement de l'Évaluation (IDEAS), coprésident du Réseau EvalIndigenous, membre d'EvalPartners et ancien président de l'Association Africaine d'Évaluation (AfrEA).

Avec plus de 20 ans d'expérience en évaluation, M. Yakeu a travaillé comme conseiller en évaluation auprès du ministère de Sécurité publique Canada; professeur invité à l'Université des Émirats arabes unis et à l'Institut Supérieur des Sciences Environnementales au Cameroun ; enseignant-conférencier auprès de l'Université Humboldt, de l'Université d'Ottawa, de l'Université du Québec à Montréal et de l'Université Constantine II en Algérie. Comme chef d'équipe, il a dirigé plus de 100 projets de

recherche et d'évaluation en Afrique, en Europe, en Amérique du Nord, au Moyen-Orient et en Asie, particulièrement pour les Nations Unies, la Banque africaine de développement, Affaires mondiales Canada, la U.S. Agency for International Development (USAID) et l'Union européenne. En 2020, il a reçu le Prix « Leadership Awards » en évaluation de « African Evidence Network » à Prétoria.

**Wend-Yam Saturnin ZOETYANDE**, M.P.H., conseiller technique suivi-évaluation et apprentissage, Pathfinder International.

M. Zoetyande est titulaire d'un master en méthodes quantitatives et économétriques pour la recherche en santé, d'une maîtrise en économie et d'une licence en statistiques sociales. Il a bénéficié de plusieurs formations en évaluation de programme via le Programme international de formation en évaluation du développement (PIFED) et le Centers for Learning on Evaluation and Results (CLEAR) du Centre africain d'études supérieures en gestion (CESAG). Il a plus de 8 ans d'expériences professionnelles dans le domaine de la recherche, du suivi et d'évaluation de projets de santé, d'éducation, etc. Il a été chargé des cours de statistiques, d'analyse des données et des méthodes quantitatives d'évaluation dans les universités publiques et privées au Burkina Faso. Il est l'actuel coprésident du réseau des jeunes et évaluateurs émergents de l'African Evaluation Association (AfrEA).