



Comment assurer le suivi et l'évaluation des agences anti-corruption : Guide à l'usage des agences, des donateurs et des évaluateurs

Jesper Johnsen, Hannes Hechler, Luís De Sousa
et Harald Mathisen

U4 est un centre de ressources en ligne qui s'adresse aux praticiens du développement souhaitant trouver des solutions aux défis liés à corruption auxquels ils peuvent être confrontés dans leur travail

Le centre de ressources U4 est géré par Chr. Michelsen Institute (CMI) qui est une fondation indépendante de recherche en sciences sociales travaillant sur les questions de développement.

U4 est financé par: AusAID (Australie), CTB/BTC (Belgique), CIDA (Canada), DFID (Royaume-Uni), GIZ (Allemagne), Norad (Norvège), Sida (Suède) and le Ministère néerlandais des Affaires étrangères.

Les points de vue exprimés dans ce rapport sont ceux du ou des auteurs et ne reflètent pas nécessairement l'opinion des agences partenaires de U4.

(Copyright 2012 - CMI/U4).

**Comment assurer le suivi et l'évaluation
des agences anti-corruption :**
Guide à l'usage des agences, des donateurs
et des évaluateurs

Jesper Johnsen
Hannes Hechler
Luís De Sousa
Harald Mathisen (coordinateur)

Traduit par Maëlle Duquesne

U4 Issue / Document d'analyse

Mai 2011 N°2

Traduction de U4 Issue 2011 N° 8



Sommaire

Acronymes	iv
Remerciements	iv
Résumé	v
1. Introduction	1
1.1 Structure du rapport.....	3
2. Les pratiques actuelles en matière d'évaluation	5
2.1 Comprendre la terminologie	6
2.2 Les résultats de l'analyse	10
2.3 De l'intérêt de produire des évaluations de meilleure qualité.....	15
3. Pourquoi le S&E est important pour les AAC	16
4. L'évaluabilité des AAC, une question complexe	19
4.1 Comprendre l'évaluabilité des AAC	19
4.2 Améliorer l'évaluabilité des AAC	24
5. Comment réaliser de meilleures évaluations ?	34
5.1 La finalité et la justification de l'évaluation	35
5.2 La composition de l'équipe d'évaluation.....	35
5.3 Le cadre, les principes généraux, la méthodologie et les questions de l'évaluation.....	36
5.4 Les méthodes d'analyse pour les évaluations	39
5.5 La participation.....	40
5.6 La garantie de qualité	41
5.7 Les aspects budgétaires et financiers du S&E.....	41
6. La méthodologie d'évaluation d'impact	43
6.1 Les composantes du modèle d'évaluation d'impact d'une AAC	44
6.2 Les pratiques actuelles en matière d'évaluation de l'impact des AAC	48
6.3 La redevabilité et l'attribution.....	49
6.4 Le principe de l'analyse contrefactuelle	51
7. Les indicateurs de suivi des performances : Introduction au catalogue d'indicateurs	53
7.1 L'élaboration d'indicateurs.....	53
7.2 Exemples d'indicateurs employés par trois AAC	55
7.3 Recommandations pour l'utilisation du catalogue d'indicateurs.....	59
8. Capitaliser les connaissances acquises dans le cadre du S&E	64
Annexe 1. Catalogue d'indicateurs	66
Bibliographie	86

Acronymes

AAC	agence anti-corruption
CAD	Comité d'aide au développement (de l'OCDE)
CdC	code de conduite
CPIA	Évaluation de la politique et des institutions nationales (de la Banque mondiale)
DFID	Ministère britannique du Développement International
ICAC	Commission indépendante contre la corruption (Hong Kong)
IMG	Indicateurs mondiaux de gouvernance (de l'Institut de la Banque mondiale)
IPC	Indice de perception de la corruption (de Transparency International)
KNAB	Bureau de prévention et de lutte contre la corruption (Lettonie)
MSI	management du système d'information
OCDE	Organisation de coopération et de développement économiques
ONG	organisation non gouvernementale
PNUD	Programme des Nations Unies pour le développement
S&E	suivi et évaluation
SFO	<i>Serious Fraud Office</i> (Bureau britannique des fraudes graves, Royaume-Uni)
TC	théorie du changement
TNA	évaluation des besoins de formation
TR	termes de référence

Remerciements

Les recherches et l'analyse présentées dans cet ouvrage ont bénéficié du soutien et de l'encouragement de l'Agence norvégienne pour la Coopération au développement (Norad).

U4 souhaite remercier Anette Simonsen (suivi et évaluation), Cathy Sunshine (rédaction et édition) et Lisa Arnestad (U4), sans lesquelles nous n'aurions pu rédiger un outil convivial, ainsi que Phil Mason (DFID), Francesca Recanatini (Banque mondiale), Anga Timilsina (PNUD) et Phil Matsheza (PNUD) pour leur précieuse contribution et leurs critiques amicales. Les auteurs assument l'entière responsabilité du contenu de ce document.

Ce rapport a été revu et validé lors d'un atelier financé par le PNUD qui s'est tenu à Bratislava les 7 et 8 juillet 2011. U4 remercie le PNUD pour cette opportunité et espère que l'organisation pourra faire bon usage de ce document dans son travail d'accompagnement des AAC à travers le monde.

En association avec :



Résumé

Le nombre d'agences anti-corruption (AAC) à travers le monde est en constante augmentation. Pourtant, donateurs internationaux et gouvernements nationaux remettent de plus en plus en question le rôle de ces organismes qui ne répondraient pas à l'attente générale.

Pour tenter de sonder la solidité des hypothèses formulées au sujet de l'efficacité des AAC, nous avons entrepris de rassembler et d'analyser les évaluations menées par des agences individuelles. Nous nous sommes malheureusement heurtés à un surprenant manque de documentation, le peu d'évaluations ayant été effectuées ne mesurant par ailleurs ni les résultats ni les impacts des AAC. Ces dernières sont exposées à de nombreuses critiques, lesquelles sont rarement justifiées par des arguments concrets. Les preuves devraient être recherchées du côté des AAC elles-mêmes. Il leur faudrait s'employer à établir des indicateurs basés sur les résultats pour démontrer la portée de leurs activités, et collecter toutes les informations nécessaires.

Pour l'heure, il est donc impossible de déterminer si l'échec prétendu des AAC relève d'un problème d'évaluation ou de conception. La valeur de ces organismes ne pourra en effet être établie que sur la base d'évaluations tangibles.

Dans cette optique, le présent document vise à fournir des informations techniques, méthodologiques et pratiques au personnel des AAC souhaitant entreprendre des activités de suivi et d'évaluation ainsi qu'à démontrer comment les résultats et l'impact du travail de ces agences peuvent être évalués à la lumière de critères factuels et objectifs.

1. Introduction

Le présent document vise à fournir des informations techniques, méthodologiques et pratiques au personnel des agences anti-corruption (AAC) souhaitant mettre en place des procédures internes de suivi et d'évaluation. Les donateurs y trouveront une méthode crédible d'évaluation des performances des AAC, et nous y décrirons également les étapes à suivre pour garantir une évaluation objective et factuelle des résultats et de l'impact du travail de ces agences. Enfin, sur la base des évaluations d'AAC que nous avons pu inventorier, nous présenterons différentes pistes d'amélioration de ces outils et verrons comment les agences peuvent consolider de manière rentable leurs systèmes et leurs procédures internes de suivi en vue de produire des informations qui leur soient réellement utiles.

Ce manuel est avant tout un guide pratique à l'intention des AAC, des donateurs et des évaluateurs, mais certaines questions stratégiques intéresseront surtout les donateurs. Du fait de l'apparition récente de la question du suivi et de l'évaluation (S&E) dans le domaine de la lutte contre la corruption, il semble en effet nécessaire d'envisager de meilleurs mécanismes de contrôle et d'évaluation pour avancer sur ce front. Mal conçues et mal mises en œuvre, les évaluations sont, au mieux, inutiles ; au pire, de la poudre aux yeux. Dans ce domaine, la qualité doit primer sur la quantité. Mesurer les performances et, surtout, l'impact des agences, n'est certes pas une tâche aisée, mais elle n'est pas impossible. Les donateurs internationaux pourraient à notre sens y contribuer, en intégrant des procédures de S&E plus poussées à leurs programmes de soutien et, par là même, en améliorant les capacités internes des AAC. Pour perfectionner leurs systèmes de S&E, ces dernières auront besoin d'une assistance technique et financière de moyen à long terme de la part des donateurs.

Malgré une réelle volonté de ces derniers de parvenir à déterminer l'impact du travail effectué par les AAC, l'examen des évaluations d'agences que nous avons effectué fait ressortir qu'ils consacrent rarement les ressources et le temps nécessaires à l'élaboration des évaluations de qualité qu'ils attendent. Contrairement aux idées reçues à ce sujet, il est impossible de mesurer l'impact des activités des AAC par le biais d'évaluations organisationnelles, car les méthodologies employées ne sont tout simplement pas conçues dans cette optique.

L'échec des AAC continue aujourd'hui de défrayer la chronique. La rhétorique autour des performances de ces organismes est passée sans transition du soutien enthousiaste au défaitisme. Après l'état de grâce de l'engagement politique des premières heures et les nombreux espoirs affichés lors du lancement en fanfare des AAC, il n'est à présent plus que question d'impasse, de manque de crédibilité et de marginalisation. En 2005, un rapport des Nations unies faisait le constat suivant :

Plusieurs pays ont créé, ou envisagent de créer, une commission ou une agence indépendante responsable de la lutte contre la corruption. La création d'une telle institution n'est toutefois pas la panacée au fléau de la corruption. On recense pour l'heure très peu d'exemples d'agences ou de commissions anti-corruption indépendantes dont la mission ait porté ses fruits. (PNUD, 2005, 5)

Naturellement, les représentants des agences individuelles réfutent systématiquement les critiques formulées au sujet des performances et de l'impact des AAC. Les experts tentent de déterminer si le prétendu fiasco de certaines agences relève d'une question d'évaluation ou de conception, la première étant plus facile à régler que la seconde. A ce stade du débat, personne n'est à même de trancher. Il ne sera possible d'effectuer une analyse éclairée de la valeur des AAC que lorsque des évaluations sérieusement documentées auront été menées – donc, que lorsque les méthodes de S&E auront été revues.

Comparés à d'autres activités gouvernementales, le suivi et l'évaluation des activités de lutte contre la corruption sont généralement négligés. L'introduction de procédures et de systèmes de base nécessite un certain investissement en ressources humaines, mais le coût d'une revalorisation significative des

méthodes de S&E – même si elles ne permettent pas une évaluation intégrale de l'impact des AAC – ne devrait pas être insurmontable. Le principal défi reste le nécessaire changement des mentalités.

La tâche confiée aux AAC n'est jamais facile ni bien définie. Ainsi, pour mener à bien leur mission de lutte contre la corruption, on attend d'elles qu'elles soient à la fois indépendantes et bien informées, qu'elles fassent preuve de compétences en matière d'application de la loi tout en jouant un rôle préventif et éducatif et qu'elles se dédient à des activités de recherches. Elles doivent composer avec les défauts des structures et des procédures traditionnelles d'application de la loi et assumer en grande partie la responsabilité de la mise en œuvre des stratégies anti-corruption nationales. Enfin, elles doivent rassurer le public sur l'engagement du gouvernement dans la lutte contre la corruption (Meagher, 2004).

La création d'une AAC peut répondre à la nécessité d'améliorer l'infrastructure éthique d'un pays, de réagir face à un scandale de corruption ou encore, de satisfaire aux exigences de traités internationaux (Meagher, 2002 ; De Sousa, 2009a, 2009b). Elle peut même constituer une manœuvre politique échafaudée par des responsables populistes cherchant à capter le discours anti-corruption (Smilov, 2009). Certaines AAC sont instituées sous la pression internationale et celle des donateurs, dans des contextes d'échec de l'Etat ou dans le cadre de programmes de renforcement des institutions nationales. Les détracteurs font valoir qu'attendre des AAC qu'elles tiennent leur rôle alors que d'autres structures de gouvernance, dont celles de soutien, sont inexistantes, sous-développées ou inopérantes, relève de la « mission impossible » et constitue un défi qu'aucune agence ne devrait être tenue de remplir (Doig, 2009). Dans ce contexte, évaluer les performances et l'impact des AAC exige de pouvoir faire la part des objectifs que les AAC sont susceptibles d'atteindre et des résultats spécifiques qui sont réellement réalisables.

Notre inventaire des évaluations d'AAC a révélé, d'une part, le nombre très limité de ces travaux et, d'autre part, leur piètre qualité. Ce problème n'est pas nouveau. Il a même été identifié il y a des années. En 2005, après avoir étudié le travail de cinq AAC africaines, Doig, Watt et Williams (2005, 5) firent le bilan suivant :

L'évaluation des performances des CAC [commissions anti-corruption], et plus particulièrement le manque d'outils de mesure appropriés et l'emploi généralisé de mesures de performances inadaptées, inutiles, irréalistes et même contre-productives, constituent un grave problème. Il devient impossible de différencier les performances organisationnelles envisageables de celles qui sont irréalisables et de distinguer les éléments que les CAC contrôlent de ceux qui leur échappent. En fait, les donateurs ne savent pas si leurs financements ont un impact sur la corruption, car ils ne le mesurent pas.

Comme nous le verrons dans ce rapport, les évaluations d'AAC reposent souvent sur des indicateurs de performance tels que le nombre de formations effectuées, le nombre de poursuites engagées, etc. Or, la plupart du temps, ces données ne sont pas systématiquement collectées, ou ne sont pas facilement accessibles. De plus, elles ne servent qu'à indiquer si une AAC a mené à bien certaines activités, et ne révèlent rien sur les résultats de ces initiatives. Certaines agences sont tout simplement dépourvues de système de suivi des performances. Les évaluations font généralement apparaître des commentaires sur les résultats et les impacts obtenus. Ces notes sont malheureusement basées sur des analyses totalement subjectives et très mal documentées et constituent plus des opinions que de réelles appréciations. En somme, la plupart des évaluations reflètent parfaitement les critiques que les parties prenantes ne connaissent déjà que trop bien. Allant à l'encontre même de l'objectif recherché, elles ne fournissent aucune recommandation claire sur les actions à venir. En outre, elles sont pour la plupart confiées à des évaluateurs et des consultants externes, parachutés quelques semaines durant dans des agences pour évaluer leurs performances. Comme nous allons le voir, ces méthodes sont problématiques à plus d'un titre.

Dans ce manuel, nous présentons une méthodologie qui vise à procurer aux AAC et aux parties prenantes extérieures (parmi lesquelles les organismes officiellement chargés de superviser les AAC, tels que la société civile et les partenaires internationaux du développement) les outils nécessaires aussi bien à la mise en œuvre d'un suivi continu des performances des AAC en matière de résultats d'activités qu'à l'optimisation de leur impact. Grâce à ces outils, les agences pourront à la fois s'assurer de répondre aux attentes de leurs nombreuses parties prenantes, regagner le soutien de la population et réduire les menaces injustifiées à leur propre existence. Si les critères minima et la méthodologie décrite ici sont correctement appliqués, les évaluations d'AAC devraient gagner en qualité et pourront enfin générer un débat éclairé sur l'efficacité de ces agences.

La méthodologie décrite a été élaborée et validée par des experts, des universitaires et des praticiens ayant tous une vaste expérience de la conception, du développement et de l'évaluation institutionnelle. Elle est principalement destinée aux AAC et à leurs partenaires nationaux et internationaux, mais pourra également être utile aux intervenants externes chargés d'évaluer le travail des agences. Nous ne cherchons pas ici à imposer une solution ou une technique d'évaluation ou de management simple et universelle. C'est précisément ce genre de simplification excessive et de démarche uniforme qui a provoqué l'échec des interventions précédentes. Chaque agence est unique et l'évaluation de ses performances requiert une approche et des indicateurs spécifiques.

La mise en place du cadre de S&E que nous proposons peut prendre un certain temps, mais l'opération est rentable sur le long terme. Un système solide de S&E axé sur les performances s'harmonise bien avec un management également axé sur les performances, et se révèle bénéfique pour la plupart des institutions publiques. Bailleurs et gouvernements se montrent de plus en plus réticents à financer des organismes aux performances incertaines et/ou auxquels le public n'adhère pas. Les AAC qui, après avoir entrepris d'améliorer leur système de S&E, enregistrent des progrès et constatent un impact croissant de leurs activités, sauront en reconnaître les avantages. Elles ne pourront mieux démontrer leur valeur à la société que si elles parviennent à rassembler des preuves crédibles de leurs performances.

Si les AAC parviennent à établir des indicateurs clairs, voire à imposer une nouvelle théorie du changement pour leurs interventions, les évaluateurs seront alors à même d'effectuer un meilleur travail. Si l'objectif des évaluations est purement comptable (et consiste à déterminer si une AAC donnée utilise bien ses ressources pour accomplir sa mission), alors les analyses superficielles habituelles suffisent. En revanche, s'il s'agit en plus de cela de tirer des leçons du travail effectué, les évaluations devront alors également cibler les résultats et les impacts obtenus. En d'autres termes, seules des évaluations complètes et réalisées sur une période plus longue permettront de déterminer les succès et les échecs des AAC et d'en déceler les causes.

1.1 Structure du rapport

Après l'introduction, le chapitre 2 s'appuie sur un inventaire des évaluations d'AAC pour passer en revue les méthodes d'évaluation pratiquées dans les agences et effectuer un tour d'horizon critique des méthodologies et indicateurs appliqués ainsi qu'une analyse des principales lacunes des évaluations traditionnelles. Nous tenterons également de clarifier la terminologie d'usage dans ce domaine. Ce chapitre s'adresse aussi bien au personnel des AAC qu'aux donateurs et aux évaluateurs.

Le chapitre 3 explique les raisons pour lesquelles il est important que les AAC développent un système de S&E et dans quelle mesure de meilleures évaluations peuvent contribuer à améliorer l'adhésion de la population et des donateurs. Ce chapitre intéressera plus particulièrement le personnel des AAC et les donateurs.

Dans le chapitre 4, nous tenterons d'apporter des éclaircissements sur l'évaluabilité des AAC et évoquerons différentes pistes pour l'améliorer. Les AAC évoluent toutes dans des conditions institutionnelles différentes et font chacune face à des problèmes particuliers dans le cadre de leur

mandat, ce qui influence leur évaluabilité. Optimiser cette dernière passe par l'introduction de théories du changement (modèles logiques), par l'instauration d'indicateurs et de systèmes de S&E, par la collecte d'informations, etc. Ce chapitre est avant tout destiné au personnel des AAC et aux donateurs.

Le chapitre 5 expose les critères et les approches méthodologiques basiques employés dans les procédures d'évaluation et propose des solutions aux lacunes identifiées dans le chapitre 2. Nous y fournissons un certain nombre d'indications sur la conception, la planification, la mise en œuvre et l'analyse critique des évaluations, à l'intention du personnel des AAC, des donateurs et des évaluateurs. De manière générale, l'approche recommandée ici est axée sur les résultats. Ce chapitre est susceptible d'intéresser un large public.

Dans le chapitre 6, nous examinerons les principes clés de la discipline particulière qu'est l'évaluation d'impact, et présenterons notamment une méthodologie à utiliser dans ce cadre. Ce chapitre s'adresse aux lecteurs intéressés par des méthodologies d'évaluation avancées.

Le chapitre 7 fournit des recommandations spécifiques sur l'élaboration d'indicateurs pour l'évaluation des AAC. Nous y présentons un catalogue d'indicateurs fonctionnels (annexe 1) ainsi qu'un certain nombre d'exemples en termes de produits, de réalisations et d'impacts. Ce chapitre devrait être utile au personnel des AAC, aux donateurs et aux évaluateurs, et plus particulièrement au personnel chargé de concevoir des indicateurs.

Pour conclure, dans le chapitre 8, nous rappellerons qu'il est important de tirer des leçons du suivi et des évaluations effectués et suggérerons différentes pistes aux AAC et aux donateurs pour revaloriser cet aspect. Les évaluateurs devraient également tirer profit de la lecture de ce chapitre.

2. Les pratiques actuelles en matière d'évaluation

En collaboration avec la Banque mondiale et le Programme des Nations Unies pour le Développement (PNUD), U4 a entrepris d'effectuer un inventaire approfondi des évaluations élaborées à ce jour, couvrant pour cela les travaux effectués par la plupart des agences anti-corruption, par les donateurs et enfin, par les consultants spécialisés dans les AAC. Cette initiative visait à rassembler le plus grand nombre possible d'évaluations d'AAC réalisées au cours des dix dernières années. La liste finale s'avéra étonnamment courte. Si, à la base, ces travaux ont fait l'objet d'une production restreinte, l'objet de notre étude s'est en plus trouvé limité du fait que consultants, AAC et donateurs aient parfois souhaité préserver une certaine confidentialité et se soient montrés inquiets de voir la diffusion publique de leurs travaux les discréditer.

Nous avons sélectionné et analysé treize évaluations, toutes de méthodologie, qualité et rigueur très disparates. Nous avons ensuite recoupé nos conclusions avec le récit des expériences des évaluateurs et des AAC. Du fait de la complexité de la réforme des institutions publiques, du caractère confidentiel des documents et du manque d'études effectuées dans ce domaine, nous avons choisi de diversifier nos sources au maximum.¹

Notre étude met en lumière le fait que nombre des « vérités » circulant au sujet des AAC ne trouvent aucun fondement solide dans les études et les évaluations systématiques existantes. Les plus fréquentes sont les évaluations de capacités organisationnelles. La plupart d'entre elles suivent – au mieux – une méthodologie d'évaluation sommaire, aucune n'appliquant de critères d'évaluation reconnus au niveau international (que nous verrons dans le chapitre 5). Par ailleurs, rares sont celles qui élargissent leurs sources au-delà des interviews avec des personnes ressources et des analyses de documents, car les services de S&E internes des AAC ne produisent généralement pas ce genre de données.

Dans ce chapitre, nous reviendrons tout d'abord sur la terminologie utilisée dans ce domaine, puis étudierons les méthodes d'évaluation des AAC, et plus particulièrement les méthodologies et les indicateurs que ces dernières emploient pour évaluer leur travail. Nous analyserons également les lacunes et les problèmes méthodologiques identifiés. L'objectif des évaluations est variable : certaines visent à améliorer la planification stratégique et les processus décisionnels de management des AAC, alors que d'autres obéissent à des impératifs de retour régulier d'informations.

La plupart des évaluations sont amorcées par des donateurs qui souhaitent effectuer un contrôle comptable et réglementaire de l'usage de l'assistance au développement allouée aux AAC, et qui mandatent pour cela des consultants. Les ressources affectées à cette tâche et les évaluateurs choisis pour les réaliser montrent que les donateurs privilégient la plupart du temps des évaluations courtes et « simplifiées » obligeant l'évaluateur à rendre un jugement subjectif des performances observées. Rares sont les évaluations entreprises en vue de démontrer que les interventions des AAC sont vecteurs de développement et, lorsque c'est le cas, le délai et les ressources impartis sont insuffisants pour que l'évaluateur puisse développer une méthodologie lui permettant de mener l'exercice jusqu'au bout.

¹ Nous avons effectué une analyse plus ou moins complète de la documentation disponible sur les AAC. Nous nous sommes également basés sur différents points de vue d'experts exprimés lors d'un e-forum hébergé par le nouveau réseau Asia-Pacific Integrity in Action (AP-INTACT). L'AP-INTACT compte plus de 120 membres, parmi lesquels des praticiens nationaux, des spécialistes internationaux en matière de lutte contre la corruption et de réforme des administrations publiques ou encore, des membres du PNUD. Le débat s'est déroulé du 16 au 27 novembre 2010.

2.1 Comprendre la terminologie

Avant de présenter notre analyse des évaluations inventoriées, nous allons tenter de clarifier les termes employés dans ce rapport. La terminologie relative au S&E utilisée dans les évaluations d'AAC nous a semblé très approximative, c'est pourquoi il est important de commencer par en définir les concepts clés.

Tout d'abord, bien que nous employions ici les termes de « projet » et de « programme » pour nous référer aux AAC, notre méthodologie rend compte du fait que ces dernières sont des organisations. Il n'est pas nécessaire de faire une distinction entre l'évaluation d'un projet à durée déterminée et celle d'une organisation (ou d'une stratégie ou encore, d'un programme). D'après Gittinger (1982), un projet se définit comme « l'ensemble des activités auxquelles est consacré un budget dans l'attente d'un retour ». Dans le cadre des évaluations, on peut donc considérer les AAC comme des projets.

Dans le domaine du S&E, la terminologie est souvent interprétée de différentes manières en fonction du public, ce qui prête à confusion. Même les termes de référence que sont ceux de « suivi » et d'« évaluation » sont ambigus. Le suivi renvoie à une procédure interne des AAC, une méthode de collecte et d'analyse d'informations que ces dernières appliquent de manière régulière et systématique en vue de déterminer si elles suivent la bonne stratégie pour atteindre leurs objectifs. Quant au terme d'évaluation, il va plus loin et peut être utilisé pour expliquer des relations de cause à effet et établir la valeur d'un projet, d'un programme ou d'une stratégie. Souvent, les spécialistes du développement emploient indifféremment les mots « appréciation » et « évaluation »². Pour les besoins de notre rapport, nous considérerons que l'appréciation d'un projet peut aussi bien inclure la notion d'appréciation préalable (appraisal) que celle d'évaluation. L'appréciation préalable d'un projet permet d'évaluer à l'avance s'il est viable ou non. L'évaluation d'un projet en mesure les performances a posteriori, c'est-à-dire après sa mise en œuvre (Université de Londres, 2010).

L'évaluation d'une AAC peut prendre différentes formes et suivre de nombreuses méthodologies. Quel que soit le type de projet en jeu – qu'il soit financé par un bailleur de fonds, un gouvernement, une organisation non gouvernementale (ONG) ou par tout autre organisme –, les instruments d'évaluation peuvent être regroupés en trois grandes catégories, en fonction du moment d'intervention dans le cycle du projet :

- *Instruction du projet ex-ante* : évalue si un projet particulier doit bénéficier d'un financement, par exemple par le biais d'une analyse coût-avantage ou d'une évaluation des risques.
- *Etude intérimaire* : évalue les performances au cours de l'intervention afin de pouvoir effectuer des ajustements au programme le cas échéant.
- *Evaluation du projet ex-post* : évalue les performances du projet après sa conclusion, par exemple par le biais de méthodologies d'évaluation qualitative ou quantitative.

Les gouvernements des pays développés exigent souvent l'utilisation de ces instruments d'évaluation dans le cadre de leurs politiques et projets, et les donateurs les emploient systématiquement (Boardman et al. 2006, 5 ; Commission européenne, 2004, chapitres 4 et 5).

² Note du traducteur : Les termes anglais d'origine, « assessment and evaluation », se traduisent tous les deux par « évaluation » en français. Pour conserver la distinction opérée dans la version originale, nous traduirons ici « assessment » par « appréciation », terme également employé dans les glossaires de l'OCDE.

Il est en revanche assez ardu d'effectuer une classification des variétés spécifiques d'instruments d'évaluation, car il n'en existe aucune liste type. Dans ce rapport, nous couvrirons cependant cinq modèles d'évaluation :

- *L'appréciation rapide (rapid assessment)*: vise essentiellement à identifier et analyser les impacts potentiels, au même titre qu'une « appréciation préalable d'impact » (« impact pre-assessment »). Utile en cas d'urgence, comme pour l'établissement de camps de réfugiés par exemple (Glasson, Therivel et Chadwick, 2005, 284–86).
- *L'évaluation des procédures (process evaluation)*: variante de l'évaluation standard de programme (voir point suivant), axée sur la mise en œuvre et sur le fonctionnement d'un projet ou d'une institution. Vise généralement à déterminer si le projet correspond à la conception initiale qui en a été faite. Les évaluations de procédures sont relativement peu coûteuses et rapides à élaborer.
- *L'évaluation de programme (programme evaluation)*: renvoie au terme « évaluation » dans son acception commune. Il existe de nombreuses sortes d'évaluations de programme, mais toutes ont pour point commun de ne pas employer de normes, d'indicateurs ou de critères prédéfinis comme le font les appréciations. Les évaluations visent à apporter des réponses en termes de conception, de mise en œuvre et de résultats. Le type de questions posées en détermine le coût, la composition et la méthode. Si ces questions sont formulées pour déterminer des relations de cause à effet, l'évaluation se base alors souvent sur un modèle logique ou sur une théorie du changement, lesquels doivent être préalablement établis.
- *L'évaluation d'impact (impact evaluation)*: évaluation spécifique basée sur une méthodologie sophistiquée utilisant l'analyse contrefactuelle pour imputer les résultats observés à l'intervention. Nous reviendrons sur cette méthode dans le chapitre 6.
- *L'appréciation des capacités organisationnelles (organisational capacity assessment)*: quel que soit le domaine, ces « appréciations de performances » (« performance assessments ») consistent généralement à comparer la réalité avec des normes et des critères prédéfinis. Ce sont des démarches similaires à des audits, qui visent à évaluer la conformité de l'intervention avec les procédures, les normes et les critères établis. Ces appréciations ne constituent pas des cadres d'évaluation. Les capacités organisationnelles désignent « l'aptitude d'une organisation à utiliser ses ressources pour fonctionner » (Lusthaus et al. 2002, 15). L'appréciation de ces capacités passe par un examen des systèmes et des procédures de gestion ainsi que des ressources humaines, financières et matérielles. Nombre d'agences de développement s'emploient à renforcer les capacités des institutions des pays partenaires, c'est pourquoi les appréciations de capacités sont devenues centrales. Il est important que ces outils ne soient pas utilisés comme substituts aux évaluations, qui ne visent pas les mêmes objectifs. Le but des appréciations de capacités est en effet de révéler les écarts entre les aptitudes souhaitées d'une institution et celles dont cette dernière dispose réellement, afin d'identifier ses besoins en soutien.³

Cette liste n'est en aucun cas exhaustive, et les modèles d'évaluation ne s'excluent pas les uns les autres. Il s'agit des principaux types d'évaluations qui peuvent être employés pour mesurer les performances d'une AAC. Soulignons également que les mécanismes internationaux ou régionaux de révision et d'autoévaluation de la lutte contre la corruption sont des outils d'évaluation de conformité (compliance review) et de mesure d'écarts (gap assessments), et non des cadres d'évaluation.

³ Les sources classiques sont Lusthaus et Anderson, ainsi que Murphy (1995) et Lusthaus et al. (2002).

Parmi les évaluations passées en revue, très peu mentionnent le modèle auquel elles se rattachent. Certaines, supposées être des évaluations de programme, adoptent plus une approche et des méthodes correspondant à des appréciations rapides ou des appréciations de capacités organisationnelles. L'ensemble s'est avéré impossible à classer de manière satisfaisante, c'est pourquoi nous avons décidé de garder le seul terme d'« évaluation » pour toutes les appréciations effectuées, sachant que peu d'entre elles concordent réellement avec cette catégorie.

Il n'existe pas de bonne ou de mauvaise sorte d'évaluation. Tout dépend de l'objectif poursuivi. Il est important de bien intégrer la corrélation entre le degré de perfectionnement de la méthodologie employée et les ressources nécessaires à l'exercice. Notre inventaire des évaluations d'AAC a en effet révélé que les évaluateurs d'agences ne disposaient souvent que de quelques semaines pour conclure leur mission, délai qui convient généralement pour une appréciation rapide. Nous avons également pu observer que certaines évaluations de résultats avaient été entreprises par le biais de méthodologies normalement appliquées aux appréciations de capacités organisationnelles, dont la portée n'est pas suffisante pour un tel exercice. Enfin, toute évaluation destinée à mesurer les résultats des AAC exige un investissement conséquent en ressources, une bonne planification et le soutien du personnel des AAC et des donateurs, à plus forte raison lorsqu'il s'agit de mesure des impacts.

ENCADRE 1. LA METHODOLOGIE DU PNUD POUR L'EVALUATION DES CAPACITES DES AAC

Le Programme des Nations Unies pour le développement (PNUD) a élaboré un « Guide à l'usage des praticiens : Evaluer les capacités des agences anti-corruption (AAC) » utilisé pour mesurer les aptitudes existantes des AAC tout en considérant les capacités nécessaires à la bonne conduite de leur mission. La méthodologie employée, qui se base donc sur celle des appréciations classiques des capacités organisationnelles, est conçue pour cibler les *fonctions* exercées par les agences plutôt que leur organisation institutionnelle ou leur nom. Cet outil peut donc tout aussi bien être utilisé pour évaluer les capacités d'une commission anti-corruption aux fonctions à la fois préventives et répressives, que d'un service préventif de lutte contre la corruption, d'un bureau d'audit se consacrant à la recherche ou encore, du bureau d'un procureur aux fonctions répressives. Il adopte un format modulaire apportant une certaine flexibilité aux AAC dans le choix d'un ensemble de fonctions clés.

Les conclusions tirées de l'évaluation des capacités d'une AAC peuvent ensuite venir alimenter un plan global de développement des capacités qui, par le biais d'un ensemble d'actions chronologiques intégrées à un programme ou à un projet, permettra de répondre aux besoins identifiés. Les indicateurs et les références établis au cours de l'analyse peuvent servir de repères pour le suivi et l'évaluation postérieurs du développement des capacités de l'agence. Au niveau national, l'évaluation doit mobiliser et impliquer tout un éventail d'acteurs de la lutte contre la corruption. Le guide fournit par ailleurs des indications détaillées sur le déroulement concret d'une procédure d'évaluation de capacités. Bien qu'il puisse être combiné avec d'autres outils, ce guide ne s'applique pas aux évaluations de résultats ou d'impacts.

Source: (PNUD, 2011)

D'autre part, il existe différents types de suivi, que l'on peut classer en fonction de leur objectif :

- *Le suivi de conformité (compliance monitoring)*: se concentre sur le respect des critères et des procédures établis.
- *Le suivi des activités (activity monitoring)* : cible la quantité, la qualité et la disponibilité des ressources et des activités.
- *Le suivi financier (financial monitoring)*: contrôle le coût de la mise en œuvre par rapport au budget et au délai prévus.

- *Le suivi organisationnel (organisational monitoring)*: cible la viabilité (sustainability), le développement institutionnel et le renforcement des capacités dans le cadre du projet et peut en signaler les points problématiques.
- *Le suivi du contexte (context monitoring)*: surveille le contexte et l'environnement du projet, et plus particulièrement les éléments susceptibles d'en affecter la mise en œuvre et la progression.
- *Le suivi des bénéficiaires (beneficiary monitoring)*: se concentre sur les impressions des bénéficiaires de l'intervention.
- *Le suivi des résultats (result monitoring)*: cible la réalisation des performances recherchées, et plus particulièrement les résultats et les impacts prévus.

Dans le cadre de leur stratégie globale de gestion, les AAC ont certainement intérêt à associer plusieurs de ces activités de suivi. En revanche, à des fins de S&E, il convient qu'elles privilégient surtout le suivi de leurs performances (qui leur permet de mesurer les résultats et les impacts de leurs activités) plutôt qu'un simple suivi de l'exécution de leur mission (correspondant au suivi cumulé de la conformité, des activités, des finances et de l'organisation en vue de mesurer leurs moyens, leurs activités et leurs produits).

Les concepts de suivi et d'évaluation mis à part, d'autres termes employés dans le vaste domaine du S&E sont parfois sujets à débat. Le mot « impact », par exemple, est aujourd'hui très répandu. Il est généralement utilisé comme un synonyme interchangeable de « résultats », « performances » ou d'autres indicateurs de rendement. Les concepts de moyens (inputs)⁴, d'activités, de produits (outputs)⁵, de réalisations/effets (outcomes) et d'impact que nous emploierons ici ont chacun une signification propre. D'après les conventions de l'Organisation de coopération et de développement économiques (OCDE), un cadre logique (souvent abrégé « logframe » en anglais) peut avoir des répercussions à trois niveaux différents, à savoir au niveau des produits, des réalisations et enfin, des impacts. Le schéma ci-dessous fait apparaître, en bleu clair, différentes significations et interprétations populaires de ces concepts.

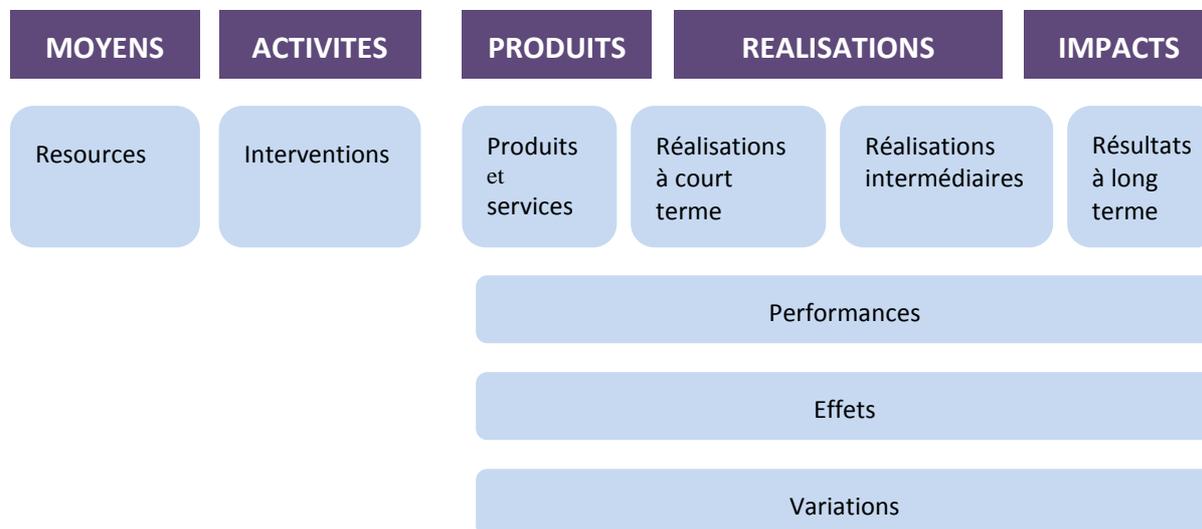
Les *moyens* sont les ressources disponibles pour un projet, un programme ou une organisation. Ils englobent les capitaux, le personnel, les installations, l'équipement et l'expertise technique. Les *activités* renvoient à l'utilisation de ces moyens, par exemple l'organisation de séminaires, l'élaboration de manuels, etc. Les *produits* sont les services ou les produits générés dans le cadre des activités d'un projet comme, par exemple, la formation de 1 500 fonctionnaires, la mise en œuvre d'un code de conduite, etc. Les *réalisations* sont les effets ou les fruits des activités et des produits. Ils reflètent l'assimilation, l'adoption ou l'usage des services et produits par les bénéficiaires supposés du projet, et révèlent les variations qui ont lieu, comme par exemple l'amélioration du niveau de confiance du public. Les *impacts* sont les conséquences à long terme d'un projet, qui renvoient généralement à la réalisation des objectifs visés. Ils sont parfois désignés comme des résultats plus généraux, comme pourrait l'être, par exemple, le « recul de la corruption ». Toutefois, comme nous le verrons plus loin dans le catalogue d'indicateurs (annexe 1), il est important, même au niveau des impacts, de spécifier quels indicateurs sont supposés rendre compte des variations. Aucun instrument ou indice de mesure de la corruption ne peut à lui seul mesurer le « niveau de corruption » en tant que

⁴ Note du traducteur : le terme anglais « inputs » est également souvent traduit par « intrants » en français. Nous emploierons ici le terme « moyens », conformément à la terminologie employée par l'OCDE.

⁵ Note du traducteur : le terme anglais « outputs » est également souvent traduit par « extrants » en français. Nous emploierons ici le terme « produits », conformément à la terminologie employée par l'OCDE.

tel. Chacun pointe plutôt différents aspects du niveau de corruption, comme la perception du public, les statistiques de criminalité, etc.

Schéma 1. Explication des concepts : des moyens aux impacts



Source : Adapté d'une diapositive présentée lors de la formation « Monitoring & Evaluation for Results », organisée par le Groupe indépendant d'évaluation de la Banque mondiale, Washington, DC, 2009.

Le terme d'« impact » est particulièrement difficile à définir. Parmi les innombrables discours sur le sujet, certains se basent sur des définitions techniques, d'autres sur l'usage courant du mot. Techniquement parlant, une évaluation d'impact est une « analyse contrefactuelle de l'impact d'une intervention sur les résultats finaux en matière de bien-être » (GIE, 2006, 1). En d'autres termes, une évaluation d'impact permettrait de mesurer le degré de réalisation de l'objectif global visé. Mais elle pourrait également permettre de mesurer les variations (impacts) produites par l'intervention, en comparaison avec le scénario qui se serait probablement déroulé si l'intervention en question n'avait pas eu lieu. Ainsi, lorsque nous décrirons plus bas les différentes méthodes de mesure d'impacts, nous ne prendrons pas en compte tous les résultats, les effets ou les variations, mais plutôt un résultat final de long terme bien précis. Comme nous l'avons vu, une intervention peut porter ses fruits aussi bien au niveau des produits que des réalisations ou encore, des impacts. Nous emploierons cependant le terme « évaluation de performances » pour décrire les procédures de S&E valorisant la mesure des résultats et des impacts.

2.2 Les résultats de l'analyse

Notre analyse se base sur un ensemble de normes de qualité élaborées par le Comité d'aide au développement de l'OCDE (CAD) en 2006,⁶ que nous avons adaptées afin de pouvoir les appliquer aux évaluations d'AAC. Nous avons ainsi obtenu les 10 critères suivants :

- Justification, finalité et objectifs de l'évaluation

⁶ La liste des critères présentée ici est basée sur les « Normes du CAD pour une évaluation de qualité » élaborées en 2006 par le Réseau du CAD/OCDE sur l'évaluation du développement, pour une application pilote de trois ans (OCDE/CAD, 2006). La première version des normes a fait l'objet d'une révision à partir de l'expérience et des apports des partenaires au développement et ont été publiées par le CAD en 2010 (OCDE/CAD, 2010b).

- Champ couvert par l'évaluation
- Contexte
- Méthodologie d'évaluation
- Sources des informations
- Indépendance
- Déontologie de l'évaluation
- Garantie de qualité
- Pertinence des résultats de l'évaluation
- Exhaustivité

Les évaluations étudiées contenant peu d'information, aucune n'a pu satisfaire à l'intégralité de ces critères. Il faut savoir que ces derniers n'apparaissent malheureusement jamais dans les termes de référence des évaluations. Nous verrons comment remédier à ce problème dans le chapitre 5.

2.2.1 Justification, objet, contexte et objectifs de l'évaluation

Presque toutes les évaluations examinées visent à évaluer l'efficacité et les performances de l'institution en jeu. Peu de cadres logiques prévoient des indicateurs de référence grâce auxquels évaluer les progrès effectués et, lorsque c'est le cas, les évaluateurs en ont souvent discuté la pertinence. De même, ces derniers ont plusieurs fois remis en cause la nécessité et l'intérêt d'un soutien des donateurs à certaines fonctions institutionnelles. Plusieurs évaluations visaient à clarifier le rôle de l'AAC en tant qu'institution, ou son rôle dans la mise en œuvre d'une stratégie nationale de lutte contre la corruption. Seule une évaluation mentionnait clairement l'évaluation d'impact comme objectif.

Sur les 13 évaluations analysées, 11 ont été commandées par les donateurs qui avaient soutenu ces mêmes AAC. Nulle part dans ces études il n'est précisé si ces derniers financent les évaluations ou s'ils en ont sélectionné les consultants. Il apparaît toutefois que certains travaux ont été entrepris au titre des examens réguliers (souvent annuels ou semestriels), à la demande des donateurs et en concertation avec les AAC. Dans certains cas, l'évaluation résulte d'un protocole d'accord entre le gouvernement et différents donateurs.

2.2.2 Champ couvert par l'évaluation

Les 13 évaluations passées en revue visent des commissions anti-corruption, des directions de lutte contre la corruption de différents ministères, des services de bureaux de procureurs généraux spécialisés dans la corruption ou encore, des groupes de travail anti-corruption.

La moitié d'entre elles omettent de mentionner la durée de l'évaluation, mais les autres évoquent un travail de 9 à 23 jours, ce qui semble extrêmement rapide étant donné la multitude de tâches assignées aux institutions et le tableau organisationnel et de gouvernance avec lequel elles doivent composer.

Aucune des évaluations ne fait référence au coût de l'exercice. Il est vrai que les financeurs peuvent préférer ne pas divulguer ce genre d'information, mais cela n'en révèle pas moins un grave manque de transparence à propos des méthodologies employées et, surtout, de leur reproductibilité. Idéalement, la publication des conditions d'évaluation et des méthodes employées devrait être la norme.

Sur les 13 évaluations, 4 ont été réalisées par un consultant seul. Deux de ces consultants étaient des ressortissants du pays des AAC, les deux autres étant des consultants internationaux également employés à titre consultatif par les institutions évaluées. Quatre évaluations ont été effectuées par des

équipes de deux personnes : dans trois d'entre elles, par des consultants internationaux, et par deux consultants locaux pour la quatrième. Les équipes plus grandes des cinq cas restants étaient généralement composées de consultants nationaux et internationaux, ces derniers étant toutefois toujours les plus nombreux.

2.2.3 Méthodologie d'évaluation

La majorité des évaluations ne suivent ou n'indiquent aucune méthodologie de recherche allant au-delà d'entretiens avec des personnes ressources et d'une analyse documentaire. On est donc en droit de remettre en question la validité, la fiabilité et la crédibilité des conclusions auxquelles elles aboutissent. Sans fondement méthodologique, les évaluations dites « rapides » se résument à de simples avis d'experts, que les décideurs politiques et les responsables de programmes ne peuvent exploiter comme renseignements fiables.

Seules deux des évaluations examinées mentionnent qu'elles sont basées sur une étude documentaire (couvrant différentes documentations, des rapports, des descriptifs de projets, des indices internationaux de corruption, etc.) et sur des entretiens avec les parties concernées, aussi bien internes qu'externes. Seules deux évaluations fondent leur analyse sur des données secondaires.

La majorité des évaluations commencent par un examen du mandat et des fonctions de l'AAC étudiée. Certaines s'intéressent également aux capacités des agences à remplir leurs fonctions, et mentionnent les ressources additionnelles (généralement, du personnel) qui leur ont été nécessaires pour mener à bien leur mission. D'autres encore évoquent parfois leur contexte organisationnel (ressources humaines et financières, systèmes d'audit et d'approvisionnement, etc.). Les questions de l'accessibilité de l'AAC et ses services, de la coordination des stratégies nationales de lutte contre la corruption et de la coopération avec la société civile sont quelquefois également abordées. De manière générale, les évaluations sont davantage axées sur les produits des AAC que sur leurs résultats.

Les évaluations évoquent presque toutes le contexte institutionnel ou juridique de l'AAC examinée, mais elles le font surtout pour expliquer l'évolution historique de l'agence. Bien que certaines rendent un avis sur l'indépendance de l'AAC ou sur sa coopération avec d'autres institutions publiques, peu d'entre elles abordent des facteurs pourtant décisifs pour la survie de cette dernière, comme l'influence politique, les mécanismes de redevabilité de l'agence elle-même ou encore, son approche des attentes du public et des décideurs. Aucune ne propose une analyse du contexte économique-politique dans lequel évolue l'AAC. Une évaluation seulement s'interroge sur le bien-fondé de l'approche institutionnelle adoptée.⁷

Les évaluations qui font appel à des indicateurs clairs pour mesurer les progrès effectués sont très peu nombreuses. Cela constitue un réel problème, dans la mesure où, si elles ne se basent pas sur les indicateurs adéquats, les conclusions des études peuvent être faussées. La plupart du temps, les évaluateurs n'ont pas le temps de les élaborer eux-mêmes. Il est donc du ressort des AAC et des donateurs de s'assurer de l'existence d'indicateurs et de systèmes de gestion de performances appropriés.⁸

⁷ L'étude de cas népalaise propose une approche intéressante associant une analyse contextuelle à une évaluation des performances. En prenant en compte les évolutions juridiques entraînant des répercussions sur la mission de l'AAC, les chercheurs ont évalué les performances de cette dernière dans plusieurs affaires pénales et observé les variations en comparant la situation avant et après la réforme législative.

⁸ Par exemple, lorsqu'en 2007, le Ministère britannique du Développement international (DFID) effectua une étude d'une AAC qu'il finançait au Sierra Leone, il observa si peu de progrès qu'il recommanda un désengagement du projet. En 2008, un examen du soutien apporté au Bureau anti-corruption du Malawi conclut

Certaines agences ont gagné en clarté dans ce domaine du fait de l'existence de cadres logiques et de plans de gestion. Mais même dans ce cas, peu d'entre elles disposent de données de référence, et les indicateurs laissent souvent une certaine marge d'interprétation. Par exemple, le « nombre d'enquêtes effectuées en un an » ne constitue pas une information à part entière, étant donné que les enquêtes peuvent durer plusieurs années et sont rarement conclues en moins d'un an. La « quantité d'enquêtes ayant abouti à des poursuites ou à un renvoi satisfaisants » pose la question de la définition du terme « poursuites satisfaisantes ». Et enfin, l'« amélioration de la qualité des enquêtes et des poursuites judiciaires » exigera certainement l'instauration de nouveaux indicateurs de suivi ou de critères de qualité plus détaillés.

En ce qui concerne les fonctions préventives, les évaluations se réfèrent généralement à des indicateurs tels que l'existence d'accords avec d'autres institutions publiques, de comités pour l'intégrité, de codes de conduite et de politiques préventives au sein des institutions publiques. La fonction de sensibilisation du public est souvent mesurée par une simple appréciation numérique des campagnes effectuées, des émissions de radio émises, des affiches imprimées, etc. Comme le font remarquer certaines évaluations, peu d'analyses portent sur l'impact en termes de compréhension de la corruption, et encore moins sur le changement des comportements.

Pour les enquêtes et les poursuites judiciaires, les indicateurs les plus utilisés sont respectivement le nombre de plaintes reçues, la part de plaintes ayant un lien avec la corruption par rapport au nombre total de plaintes, le nombre d'enquêtes effectuées, la durée des enquêtes, les affaires classées ou ayant donné lieu à des poursuites, le nombre de poursuites engagées et enfin, le nombre d'acquittements prononcés. Ces indicateurs ciblent tous les activités entreprises par les AAC, sans presque aucune référence à leur résultat ou à leur impact éventuel sur la corruption dans une zone à risque donnée. Certains indicateurs employés ciblent davantage les réalisations, et mesurent le taux de condamnations et le recouvrement d'avoirs.

Il apparaît clairement que les évaluations auraient pu mieux valoriser les indicateurs et souligner l'importance de disposer de cadres logiques bien pensés. Il semble que les donateurs et le personnel des AAC aient élaboré des indicateurs non pas pour les besoins de l'évaluation, mais plutôt dans une optique de gestion de programmes. L'évaluation s'appuie donc sur une base plutôt fragile. Nous suggérerons plus loin de quelle manière les cadres logiques peuvent être conçus de manière plus collaborative avec les AAC et comment ces dernières peuvent les employer, non seulement dans les évaluations, mais également dans leurs activités quotidiennes en tant qu'instruments de gestion axée sur les performances.

2.2.4 Sources des informations, indépendance, déontologie de l'évaluation et garantie de qualité

Globalement, les sources utilisées dans le cadre des évaluations sont très mal référencées, ce que l'on peut considérer comme un manquement à des normes d'évaluation pourtant fondamentales. Cependant, une lecture attentive fait clairement apparaître que, dans la majorité des cas, les évaluations se basent sur des analyses documentaires et des entretiens avec des personnes ressources.

Malheureusement, les évaluations qui fournissent la liste des documents consultés et des personnes interrogées sont minoritaires. Celles qui ne donnent pas les sources des données utilisées manquent de crédibilité. De leur côté, les évaluations qui mentionnent ces informations font référence à tout un éventail de parties prenantes interviewées, dont le nombre et la portée ont apparemment été limités par le manque de temps.

à des progrès satisfaisants, mais qui avaient fait l'objet d'une très mauvaise communication. En 2010, un examen annuel de la même institution détermina que les objectifs avaient été dépassés et devaient être réajustés.

On peut donc s'interroger sur l'indépendance et l'objectivité des évaluations. Constituent-elles un compte-rendu honnête de la situation ? Les évaluateurs adoptent-ils un point de vue positif pour satisfaire leur client, et négatif lorsqu'ils pensent que c'est ce que l'on attend d'eux (par exemple en cas d'intention d'interruption des financements) ? Autrement dit, sont-ils objectifs ou partiaux ?

L'étude des évaluations ne permet pas de déterminer directement celles qui sont financées, et donc potentiellement influencées, par les donateurs. Toutefois, les entretiens réalisés dans le cadre de notre étude semblent indiquer que les donateurs exercent souvent une certaine influence sur le résultat des évaluations. D'après certains évaluateurs expérimentés et anciens hauts fonctionnaires, les donateurs obtiendraient toujours les conclusions qu'ils souhaitent, qu'elles soient positives ou négatives. Bien que cela relève avant tout d'une question politique, l'emploi de méthodologies d'évaluation plus rigoureuses pourrait limiter la subjectivité et l'arbitraire de l'exercice et donc réduire les risques de le voir aboutir à des conclusions faussées. L'emploi de critères bien définis et l'application de procédures de contrôle de la qualité des évaluations pourraient par exemple y contribuer. Nous y reviendrons dans le chapitre 5.

En ce qui concerne les garanties de qualité, on peut penser que la majorité des évaluations financées par des donateurs ont été partagées de manière *ad hoc* avec les agences. Permettre à l'institution examinée de commenter les résultats auxquels a abouti l'évaluation n'est pas seulement une question de transparence et de respect de normes éthiques ; c'est également un moyen de réduire les marges d'erreur, de produire des données de meilleure qualité et de valider les conclusions tirées. D'après les évaluations que nous avons passées en revue, il semble toutefois que ce genre de pratique soit peu fréquent. Donateurs et AAC devraient s'employer ensemble à promouvoir des procédures de contrôle de la qualité de leurs évaluations mieux conçues et plus transparentes. Nous verrons les différentes possibilités qui s'offrent à eux dans ce domaine dans le chapitre 5.

Très peu d'évaluations précisent si la procédure a impliqué un service de S&E interne à l'AAC. Que de tels services soient inexistantes ou inefficaces, il est de toute façon frappant de constater que les recommandations n'évoquent pas le besoin de les renforcer et de les associer aux futures évaluations. C'est d'autant plus surprenant que les services de S&E devraient jouer un rôle majeur dans la production de données sur les indicateurs clés et dans le suivi des progrès effectués entre les évaluations. La plupart du temps, ils pourraient également être les fers de lance des évaluations de performances et de résultats et/ou d'impact, ainsi que de la communication.

2.2.5 Pertinence des résultats de l'évaluation

Ce critère est sans doute le plus important des dix, puisqu'il se concentre sur la valeur initiale de l'évaluation pour ses parties prenantes et ses bénéficiaires. Si les AAC souhaitent réellement tirer profit des évaluations, elles doivent alors se focaliser sur les leçons qu'elles peuvent en tirer (voir chapitre 8 pour une analyse plus approfondie de cette question).

Nous nous félicitons de constater que certaines évaluations s'interrogent sur les suites qui ont pu être données aux recommandations formulées dans les précédentes évaluations. Les comptes-rendus d'entrevues semblent cependant indiquer que les évaluations commanditées de l'extérieur sont transmises à la direction mais ne font pas l'objet d'une diffusion plus étendue dans les agences. On peut donc penser que les AAC utilisent davantage les évaluations pour rendre des comptes aux institutions extérieures que comme outils de capitalisation de leur expérience. Une agence qui souhaite exploiter pleinement une évaluation se doit d'en assurer une large diffusion. Déterminer dans quelle mesure le personnel des AAC se soucie de la pertinence des évaluations et à quel point il exploite les conclusions qui en sont tirées n'entre toutefois pas dans le cadre de notre rapport.

2.3 De l'intérêt de produire des évaluations de meilleure qualité

A la lumière de l'analyse sommaire des méthodes d'évaluation des AAC, nous sommes convaincus que les méthodologies employées par la communauté des donateurs peuvent être améliorées. En perfectionnant les évaluations des résultats et de l'impact des activités des AAC, les donateurs disposeraient d'une base solidement documentée sur laquelle fonder leurs décisions de renforcer leur soutien aux agences ou au contraire de s'en désengager. Un tel progrès serait également utile aux évaluateurs, qui ne suivent à l'heure actuelle aucune méthodologie rigoureuse. Les AAC en seraient les premiers bénéficiaires : elles renforceraient leur propre système opérationnel, seraient plus à même de capitaliser leur expérience et, en définitive, seraient plus efficaces dans la lutte contre la corruption.

3. Pourquoi le S&E est important pour les AAC

Tout organisme public devrait être tenu de rendre des comptes sur l'utilisation qu'il fait des ressources publiques. Pour cela, un système d'évaluation des performances doit être mis en place. Du fait de leur mandat unique, les AAC devraient également être en mesure de démontrer leur efficacité et l'impact de leurs activités aux citoyens des pays dans lesquels elles interviennent.

Les belles déclarations et autres « histoires » sur l'efficacité des AAC ne sont la plupart du temps pas corroborées par des données factuelles. Pour faire entendre leur cause dans un environnement de plus en plus concurrentiel pour les donateurs et les financements nationaux, et dans lequel les initiatives pour la gouvernance sont souvent remises en question, les AAC doivent s'engager à effectuer un suivi et une évaluation de manière sérieuse. Pour ce faire, elles doivent dépasser leurs pratiques actuelles et étayer le fruit de leur travail de preuves tangibles.

Avant toute chose, il convient de bien définir le terme de S&E. Le *suivi* se réfère au « processus continu de collecte systématique d'informations, selon des indicateurs choisis, pour fournir aux gestionnaires et aux parties prenantes d'une action de développement en cours, des éléments sur les progrès réalisés, les objectifs atteints et l'utilisation des fonds alloués. » (OCDE/CAD, 2002, 27). L'*évaluation* désigne « l'appréciation systématique et objective d'un projet, d'un programme ou d'une politique, en cours ou terminé, de sa conception, de sa mise en œuvre et de ses résultats. Le but est de déterminer la pertinence et l'accomplissement des objectifs, l'efficacité en matière de développement, l'efficacité, l'impact et la durabilité. 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. » (OCDE/CAD, 2002, 21). Les exercices de suivi et d'évaluation sont donc liés mais indépendants, et suivent des procédés et des objectifs différents.

Au vu des investissements effectués dans les AAC, on pourrait penser que le S&E fait partie intégrante de leurs activités. Les conclusions de notre étude indiquent le contraire. Pour la plupart, les AAC devraient donc prendre un peu de recul et revenir à l'essentiel avant d'envisager de perfectionner leurs activités d'évaluation. Il est fondamental qu'elles instaurent un cadre de S&E dès le départ et qu'elles en définissent les cibles et les délais à l'avance. Elles devraient prendre le temps d'élaborer de bons indicateurs, y compris au niveau des résultats et des impacts, et qu'elles apprennent à diffuser leurs résultats (De Sousa, 2009a, 2009b).

Les cibles devraient être définies par les AAC en concertation avec le gouvernement, les donateurs, les communautés et les organisations de la société civile, de manière à en assurer l'appropriation et à instaurer une redevabilité mutuelle des parties concernées. Aucun changement inopiné des cibles – c'est-à-dire, des objectifs – ne devrait être autorisé.

Les organisations se montrent souvent réticentes à entreprendre le suivi et l'évaluation de leurs performances, qu'elles considèrent comme un exercice onéreux et fastidieux qui entrave leur « vrai travail » quotidien. Le S&E a deux grandes finalités : la redevabilité et la capitalisation de l'expérience. L'obligation de rendre des comptes est généralement ce qui pousse les AAC à entreprendre des procédures de S&E pour les projets financés par des bailleurs de fonds. Ces

ENCADRE 2. POURQUOI EST-IL IMPORTANT D'ÉVALUER LES RESULTATS ?

- Sans évaluation des résultats, il est impossible de distinguer le succès de l'échec.
- S'il est impossible de constater le succès, il est impossible de le récompenser.
- S'il est impossible de récompenser le succès, c'est probablement l'échec qui est récompensé.
- S'il est impossible de constater le succès, il est impossible d'en tirer des leçons.
- Si l'échec n'est pas reconnu, il ne peut être corrigé.
- S'il est possible de démontrer des résultats, il est possible de gagner l'adhésion du public.

Source: Osborne et Gaebler, 1992.

derniers sont en effet soucieux de savoir comment leurs fonds sont dépensés et s'ils le sont à bon escient. Mais le S&E est également un outil de capitalisation d'expérience. Comme nous l'avons vu, les systèmes de gestion axés sur les performances sont basés sur les mêmes principes de collecte d'informations sur les activités menées, qui, à leur tour, contribuent à l'amélioration du « vrai travail » des AAC.

Les procédures de S&E permettent aux AAC de mettre en avant les bénéfices, les risques et les limites de leurs activités. Elles ne constituent pas seulement un impératif destiné à faire valoir les fruits d'un projet auprès des donateurs et des contribuables, mais bien un processus de capitalisation d'expérience qui peut induire un changement stratégique et reflète les impacts des activités des AAC dans leur contexte.

ENCADRE 3. POURQUOI LE S&E EST-IL IMPORTANT AU REGARD DES ACTIVITES ET DE L'EXISTENCE DES AAC ?

Transparence : S'assurer une bonne visibilité et entreprendre des actions directes auprès de la population sont des démarches importantes pour gagner le soutien du public. Une AAC qui reste inconnue du grand public est condamnée à la marginalisation, voire à la disparition. Les procédures de S&E génèrent des rapports écrits et des chiffres concrets sur les résultats des agences, contribuant de fait à leur transparence et à leur visibilité.

Redevabilité : Les AAC sont financées par des fonds publics. Dès lors, il leur faut rendre des comptes sur leurs activités, leurs difficultés et leurs résultats auprès de ceux qui les financent, à savoir, les contribuables et les donateurs. Le S&E est une source fiable d'informations sur les performances des agences, qui favorise un suivi simple et systématique des progrès réalisés.

Mémoire institutionnelle : Les exercices de S&E permettent aux gouvernements et aux donateurs de suivre les performances et l'évolution des capacités d'une AAC sur une certaine période, et donc de déterminer si elle est apte à remplir sa mission. Les agences peuvent en outre développer une mémoire interne sur leurs différentes phases d'institutionnalisation.

Capitalisation : Le S&E produit des données tangibles à partir desquelles les AAC peuvent reconsidérer et tester certains postulats et intégrer les leçons tirées et l'expérience acquise à leurs stratégies et pratiques. Il constitue à la fois un point de départ et une procédure de réflexion. Enfin, il fournit des outils plus consistants aux AAC dans leur recherche de financements et leur orientation stratégique.

Valorisation stratégique : Des cadres de S&E solides permettent aux dirigeants des AAC et des gouvernements de déterminer si les stratégies établies fonctionnent comme prévu, en détectant les risques et les problèmes opérationnels. D'où viennent les problèmes ? Dans quelle mesure les performances de l'agence sont-elles affectées ? De quelles capacités / ressources dispose-t-on pour réduire ces risques et problèmes, et peuvent-elles être renforcées ?

Amélioration des résultats : Chacun des points mentionnés ci-dessus devrait contribuer à améliorer les résultats des AAC dans leurs activités de lutte contre la corruption.

Les rapports de S&E révèlent souvent des erreurs commises lors de la réalisation d'études et de recherches. Evaluer les méthodes de promotion d'une campagne de sensibilisation du public ou de mise en œuvre d'un système de traitement des plaintes peut fournir une occasion de remettre en question certains postulats de base et de s'interroger sur « les pratiques habituelles en vigueur », et donc d'innover et de se rénover. Les exercices de S&E conçus comme des outils de capitalisation d'expérience constituent donc également d'importantes activités de renforcement des capacités, qui peuvent déboucher sur des routines et des activités utiles aux AAC et les motiver à entreprendre des procédures de S&E, pour en tirer des leçons.

L'instauration de services de S&E internes est bénéfique à plusieurs titres, du fait que cela favorise la redevabilité et la capitalisation de l'expérience. En revanche, limiter les évaluations à des exercices externes confiés à des experts internationaux « parachutés » quelques jours dans les AAC court-circuite complètement toute possibilité de capitalisation des données.

4. L'évaluabilité des AAC, une question complexe

Ce chapitre sert deux objectifs. Tout d'abord, la première partie vise à clarifier le concept d'« évaluabilité » des AAC. Une fois cela établi, nous proposerons un certain nombre de stratégies pour l'amélioration de l'évaluabilité des agences.

Dans le domaine de la lutte contre la corruption, le S&E est largement sous-exploité. Cela tient peut-être au peu de considération accordée à l'importance de preuves probantes et aux outils de mesure, et également au manque de compétences techniques dans ce domaine pour entreprendre des évaluations « professionnelles ». Peut-être encore est-ce simplement le symptôme de cette idée répandue qui veut que le domaine de la lutte contre la corruption comporte des difficultés « particulières » qui rendraient les évaluations rigoureuses impossibles et futiles.

Dans ce chapitre, nous présenterons les difficultés de terrain les plus fréquentes, puis nous détaillerons les caractéristiques institutionnelles que les évaluateurs doivent prendre en compte et enfin, nous ferons une présentation rapide des méthodes qui peuvent contribuer à améliorer l'évaluabilité des AAC. Nous concluons sur la thèse que les AAC ne sont pas faciles à évaluer, mais qu'elles sont « évaluables » si elles développent certains éléments en interne. Aboutir à des analyses crédibles exige l'élaboration et l'utilisation d'une approche méthodologique rigoureuse, d'une chaîne de résultats et d'indicateurs, et passe par la collecte et l'analyse minutieuse des données.

Le fait que les AAC soient des organisations polyvalentes poursuivant des objectifs pluriels dans un domaine où les données ne sont pas toujours facilement accessibles ne signifie pas qu'elles ne puissent être évaluées que par le biais de ce que l'on appelle les évaluations de capacités organisationnelles. Des organisations polyvalentes d'approche tout aussi complexe, comme les Fonds sociaux pour le développement,⁹ existent dans d'autres secteurs, dont le travail sur le développement des capacités internes de S&E a contribué à améliorer l'évaluabilité.

Le présent chapitre s'adresse aux trois groupes cibles de ce rapport : le personnel des AAC, les donateurs et les évaluateurs. Les membres des AAC y apprendront quels éléments de leur système interne il convient de renforcer avant d'envisager une évaluation approfondie de leur travail. Les donateurs y trouveront le détail des critères qui déterminent le moment opportun pour entreprendre l'évaluation d'une AAC et des suggestions pour améliorer les évaluations. Enfin, les évaluateurs y obtiendront les informations nécessaires sur les caractéristiques institutionnelles spécifiques des AAC, sur les difficultés particulières auxquelles ils peuvent avoir à faire face dans ce domaine et sur les méthodes à leur disposition pour surmonter ces obstacles.

4.1 Comprendre l'évaluabilité des AAC

Une grande partie des évaluations inventoriées pour les besoins de ce rapport n'auraient pas dû être entreprises, pour la simple raison qu'elles ne s'appuient que sur très peu de données tangibles et restent totalement subjectives. Les évaluateurs n'en sont pas les responsables. C'est plutôt le reflet d'une certaine ignorance de la nécessité de mettre en place différents paramètres fondamentaux avant que les évaluateurs ne puissent entamer leur travail. En d'autres termes, avant de mandater un évaluateur, il conviendrait de s'assurer de l'existence d'objectifs clairs, d'indicateurs de résultats et de systèmes de S&E, pour ne citer que quelques exemples.

⁹ L'évaluation d'impact du Fonds social pour le développement au Yémen en constitue un bon exemple (Recovery & Development Consortium, 2010).

Le concept d'« évaluabilité » renvoie au degré de fiabilité et de crédibilité d'évaluation d'une activité ou d'un programme donné. L'appréciation de l'évaluabilité suppose d'examiner à l'avance l'activité projetée afin de vérifier si ses objectifs sont définis de façon adéquate et si les résultats sont vérifiables (OCDE/CAD, 2000, 2009). Cela permet en outre de révéler d'éventuelles difficultés susceptibles d'empêcher l'élaboration d'une évaluation utile. L'appréciation de l'évaluabilité « requiert d'examiner la cohérence et la logique du programme concerné, de déterminer la disponibilité des données, de définir l'usage potentiel que les dirigeants et les parties prenantes feront des conclusions de l'évaluation en fonction de leurs intérêts et de planifier toute évaluation par rapport aux programmes ou décisions politiques à venir » (Commission européenne, 2009).

Lorsqu'un projet n'est pas prêt à être évalué, il convient de formuler des recommandations en vue de le préparer à une telle procédure. Les défauts les plus fréquents se rapportent à trois grands domaines : la conception des programmes, la disponibilité des informations pertinentes et le choix du moment.¹⁰ Des guides et check-lists existent qui proposent des conseils pour mesurer l'évaluabilité d'un projet.¹¹

4.1.1 Les principales difficultés d'appréciation sur le terrain

Fredrik Galtung faisait observer à juste titre que « jusqu'à la moitié des années 1990, dans la documentation universitaire, les analyses empiriques effectuées dans le domaine de la corruption étaient, pour la plupart, secondaires ou anecdotiques » (2005, 101). Dans un état des lieux de la recherche sur la corruption qu'il rédigea en 1997, Bardhan déplora sincèrement que « [son] approche dans ce document [soit] avant tout analytique et spéculative, en raison des difficultés inhérentes de collecte (et donc de l'inexistence) de données empiriques fiables sur le thème de la corruption » (1997, 1320).

Les chercheurs en sciences sociales qui souhaitent étudier la corruption doivent faire face à deux grandes questions :

Comment étudier une chose en partie définie par le fait que des individus s'emploient par tous les moyens à la dissimuler ? Comment mesurer avec précision l'ampleur de la corruption quand l'intention même de la mesurer peut inciter les acteurs en jeu à limiter leur comportement illicite le temps nécessaire ou à chercher de nouveaux moyens de se cacher ? (Banerjee, Hanna et Mullainathan, 2009, 1)

Les praticiens anti-corruption s'accordent généralement à penser que « la science de la mesure de la corruption est plus une forme d'art qu'un processus empirique bien défini » (PNUD, 2008, 3). Évaluer la lutte contre la corruption est une gageure du fait de la difficulté d'élaborer des outils, des indicateurs et des statistiques fiables sur le phénomène souterrain, illégal et fortement lié au contexte qu'est la corruption (Langseth, 2006, 16).

Cette liste de doléances n'est pas l'apanage du domaine de la lutte contre la corruption : nombreuses sont les sphères qui se considèrent comme « particulièrement complexes ». Néanmoins, il existe aujourd'hui tout un éventail de méthodes grâce auxquelles les évaluateurs peuvent produire des informations et des données tangibles sur la lutte contre la corruption. Nous les décrivons dans le chapitre 5.4.

La question qui se pose toutefois est de savoir s'il est vraiment nécessaire de mesurer le niveau de corruption pour définir les performances d'une AAC. Pour commencer, il conviendrait de déterminer si le niveau de corruption constitue bien la variable clé dans l'étude des performances des agences. La

¹⁰ L'UNIFEM (2009) propose une check-list utile en pages 3 et 4.

¹¹ Consulter, par exemple, Wholey, Hatry et Newcomer (2004, 33–41).

plupart d'entre elles se consacrent probablement davantage à coordonner des activités stratégiques qu'à lutter contre la corruption. Lorsqu'elles se voient confier des missions directes de lutte contre la corruption, elles ciblent généralement un genre de corruption spécifique, comme la grande corruption, la corruption politique ou la corruption du secteur public.

Il est intéressant de noter que les critères d'évaluation de l'OCDE utilisés dans le cadre des évaluations routinières de programmes ne renvoient pas nécessairement au niveau de corruption. L'« efficacité » et l'« efficience » d'une AAC sont en effet évaluées d'après les procédures, les produits et les ressources de l'agence, qui sont comparés à ceux d'autres organisations similaires – autant de mesures qui ne dépendent pas de l'estimation des niveaux de corruption.

Le critère d'« impact » est plus problématique. Pour comprendre l'impact d'une AAC, on mesure généralement les variations du niveau de corruption (le type de variation ciblé peut varier en fonction du mandat, des activités et des domaines d'intervention des AAC). Toute agence publique suit des objectifs généraux ; pour interpréter leurs performances, il faut évaluer les variations relatives à ces objectifs. La tendance alors est de se replier sur des macro-indicateurs génériques tels que l'Indice de perception de la corruption (IPC) de Transparency International, qui apparaît comme la seule source de données disponible en matière d'impact. Malheureusement, ces indicateurs sont si généraux qu'ils saisissent rarement la nature exacte de l'impact souhaité d'une AAC.

Prenons l'exemple d'un autre secteur : l'évaluation d'un programme sanitaire de vaccination d'enfants ne consiste pas uniquement à mesurer le taux de longévité (indicateur d'impact global dans le domaine de la santé). Beaucoup considèrent que d'autres aspects de la santé sont tout aussi importants, sinon plus, comme la qualité de vie, l'absence de maladies graves, etc. En outre, l'impact poursuivi dans le cadre de ce programme cible uniquement des enfants. La définition du concept de santé varie en fonction des activités évaluées ; il en va de même pour celui de lutte contre la corruption. Les indices nationaux de perception ou de contrôle de la corruption constituent certes des informations générales, mais ils restent incomplets et ne devraient pas être utilisés à la place des analyses effectuées au niveau des projets ou des organisations.

Serait-il alors impossible d'évaluer l'impact des activités de lutte contre la corruption ? La réponse est non. La tâche nécessite toutefois l'identification préalable des variables spécifiques de produits, de réalisations et d'impact à mesurer. La question des impacts exige que les évaluateurs fassent preuve de sens commun. Faire reculer le niveau de corruption dans un pays donné est une entreprise de taille, qui requiert des investissements considérables en ressources et en moyens de la part de la multitude d'acteurs impliqués au niveau national dans la lutte pour l'intégrité. C'est une tâche tout aussi vaste que l'amélioration de la santé publique ou la réduction de la pauvreté dans un pays. Une évaluation pourrait-elle rendre un programme ou une organisation sanitaire ou de lutte contre la pauvreté responsable du niveau général de santé et de pauvreté dans un pays ? Cela peut arriver, mais seulement si le programme en question emploie d'importantes ressources et que les mesures se font sur le très long terme. Pour les programmes plus modestes menés sur le moyen terme, il est plus productif de décomposer les concepts généraux tels que la corruption en indicateurs plus précis avant d'entreprendre une évaluation.

4.1.2 Les différentes caractéristiques institutionnelles des AAC

Le monde des AAC est bien plus atypique et hétérogène que ce que l'on pourrait croire. Il n'est donc pas recommandé d'appliquer indifféremment les mêmes méthodologies et « boîtes à outils » standardisés à toutes les agences. La conception d'une AAC ne devrait normalement pas influencer son évaluabilité, mais les évaluateurs doivent être conscients des différences potentielles de logique d'intervention et de théories du changement d'une agence à l'autre. Les évaluations doivent par conséquent intégrer le contexte et les difficultés particulières à chaque AAC. Certains pays décident de mettre en place de nouvelles institutions de lutte contre la corruption ; d'autres changent simplement le nom ou les compétences d'organismes existants ; d'autres encore préfèrent créer des AAC au rôle à la

fois préventif, éducatif et informatif. Certains vont même plus loin et instaurent des AAC polyvalentes spécialisées dans la prévention, la détection et, parfois, la répression de la corruption.

La documentation sur le sujet identifie toute une série de conditions pour qu'une agence soit définie comme une AAC (De Speville, 2000 ; Doig, 1995, 2000 ; Pope et Vogl, 2000 ; Pope, 1999 ; Quah, 2000). Dans la pratique, toutefois, très peu d'AAC remplissent l'intégralité de ces critères. Il n'existe aucun modèle standardisé d'AAC. Certaines sont créées en tant qu'organismes publics ; beaucoup sont constituées en tant que services spéciaux au sein de bureaux de médiateurs et de procureurs, d'agences nationales d'audit ou encore, de forces de police ; d'autres encore ont des mandats purement préventifs ou de coordination – un rôle limité qui pourrait laisser à penser que la création de ces institutions vise uniquement à satisfaire la demande des donateurs.

Dans un rapport intitulé *Specialized Anti-Corruption Institutions: Review of Models* (OCDE, 2007), l'OCDE tente de dresser un classement des agences en trois grands types :

- *Les institutions spécialisées aux compétences multiples* : Désigne les agences anti-corruption *stricto sensu*, qui exercent des pouvoirs préventifs et répressifs et mènent toute une gamme d'activités allant au-delà des enquêtes criminelles traditionnelles, comme l'analyse stratégique, le conseil et l'assistance technique, la diffusion d'informations, la veille éthique, la formation et la recherche scientifique (évaluations des risques, études, etc.).¹²
- *Les services spécialisés au sein de forces de police ou les services en charge des poursuites*: Un même organisme cumule parfois des compétences en matière d'enquête et de poursuites judiciaires, comme c'est par exemple le cas de la Direction nationale de lutte contre la corruption en Roumanie. Il est difficile de faire une distinction claire entre ce type d'agence et les deux autres, car certaines exercent également des fonctions préventives.¹³
- *Les institutions aux compétences exclusivement préventives* : Ces organismes peuvent effectuer des recherches scientifiques dans le domaine de la corruption, élaborer des stratégies de contrôle de la corruption et conseiller les organes décisionnaires, effectuer une veille de la réglementation dans les secteurs à hauts risques, aussi bien publics que privés (par exemple, l'urbanisme, les travaux publics, les marchés publics, les casinos, les autorités douanières, etc.) et proposer des modifications en la matière, assurer une surveillance des règles en matière de conflits d'intérêts et de déclarations de patrimoine, rédiger des codes de conduite et les mettre en œuvre, assister les fonctionnaires sur les questions de corruption, faciliter la coopération internationale dans ce domaine ou encore, faire office d'intermédiaires entre la société civile et les organismes publics compétents en matière de corruption.¹⁴

¹² Parmi les agences de ce type, on peut citer par exemple la Commission indépendante contre la corruption de Hong Kong, le Bureau d'Enquête sur les pratiques de corruption de Singapour, la Commission indépendante contre la corruption de Nouvelle-Galles du Sud (Australie), la Direction de la corruption et de la criminalité économique du Botswana, l'Unité d'enquête spéciale de République de Lituanie et enfin, le Bureau de prévention et de lutte contre la corruption de Lettonie (KNAB).

¹³ Ou encore, l'Autorité nationale d'enquête et de poursuites des crimes économiques et environnementaux, en Norvège, le Bureau central pour la répression de la corruption en Belgique, la Fiscalía Anticorrupción en Espagne, la Direction centrale d'enquête sur la corruption et les crimes économiques et financiers de la Police judiciaire au Portugal et le Service central pour les poursuites pénales et l'investigation du Bureau du Procureur général au Portugal.

¹⁴ Par exemple, le Service central de prévention de la corruption (SCPC) en France, la Commission d'Etat pour la prévention de la corruption en Macédoine, le Groupe de surveillance contre la corruption en Albanie, la Commission permanente contre la corruption à Malte et le Bureau d'éthique gouvernementale aux Etats-Unis.

Les attentes à l'égard des AAC devraient être à la mesure des pouvoirs et des capacités de ces dernières. Cela peut sembler évident, mais la question est souvent négligée. Par exemple, une AAC aux compétences purement préventives ne devrait pas être évaluée à la lumière d'indicateurs relatifs à l'application de la loi. Nous développerons davantage ce point dans les chapitres 5 et 6.

Il conviendrait que les évaluateurs assimilent un certain nombre de caractéristiques institutionnelles et de paramètres externes qui ont généralement une grande influence sur les performances des AAC. Les facteurs susceptibles d'entraver la mission des agences à différents niveaux sont les suivants :

- *Les facteurs contextuels* : Le succès des AAC doit être envisagé dans son contexte. Les agences évoluent dans un environnement déterminé par différentes cultures organisationnelles, ainsi que par le niveau de développement national et par la stabilité politique du pays. On ne peut par exemple pas attendre d'une AAC qu'elle fonctionne correctement dans un pays connaissant de graves problèmes de gouvernance (De Speville, 2000 ; Doig, Watt et Williams, 2007).
- *Le cadre juridique et les facteurs politiques* : Pour mener à bien leur mission, les AAC doivent composer avec les choix politiques des élus, qui peuvent déterminer leur statut juridique, leur emplacement, leurs pouvoirs spéciaux, le degré de partage des compétences et des informations, leur autonomie financière, leurs obligations de rendre des comptes sur leurs activités, etc. De par leur statut, les AAC sont en quelques sortes toutes « indépendantes ». Dans la pratique, toutefois, leur degré d'autonomie opérationnelle varie considérablement d'une agence à une autre et, bien souvent, ces organismes ne sont opérationnellement et financièrement indépendants que sur le papier.
- *Les facteurs organisationnels* : Les mauvaises performances des AAC peuvent également être imputables à l'inadéquation de leurs procédures de recrutement et de redevabilité, ou encore à l'inadaptation ou à l'inexistence de leurs mécanismes de gestion, pourtant indispensables à leur bon fonctionnement. Les capacités techniques des AAC peuvent également pâtir d'une collaboration inefficace avec les différentes autorités compétentes. Une agence qui peine à obtenir des preuves tangibles des pratiques de corruption ou des informations sur les zones à risques de la part des organismes publics ne peut être complètement efficace. En outre, les AAC ne jouissent que rarement de l'autorité nécessaire pour s'assurer que les institutions publiques appliquent bien les recommandations qu'elles leur ont adressées. Le cadre juridique régissant la collaboration interinstitutionnelle n'est généralement pas clairement défini au départ. De plus, les AAC sont souvent sollicitées à la limite de leurs possibilités pour remplir des mandats bien trop vastes. Les agences polyvalentes doivent jongler avec le fait qu'on attende d'elles qu'elles stoppent la corruption (approche fondée sur la peur) tout en exerçant une fonction de conseil en matière de prévention (approche fondée sur la confiance).
- *Les facteurs financiers* : Toute organisation vit dans la crainte permanente de manquer de ressources financières. Les AAC disposant de gros budgets ne sont certes pas nécessairement les plus productives ; il n'en reste pas moins que celles qui opèrent dans des conditions précaires voient d'emblée leurs objectifs sérieusement compromis.
- *Les facteurs d'expertise et de leadership* : Les compétences, l'expérience et les connaissances individuelles des membres du personnel des AAC sont des éléments déterminants dans le succès de ces dernières. Les capacités des agences sont susceptibles d'être affectées tant sur le plan technique qu'au niveau de leur fonctionnement général (qui implique notamment le leadership, la gestion des ressources humaines, la planification ou encore, l'apprentissage organisationnel). Les AAC ont cet avantage sur les organismes traditionnels de détection et de répression qu'elles sont capables de générer une approche de la lutte contre la corruption fondée sur l'exploitation des connaissances acquises par le biais d'évaluations de risques et d'autres études additionnelles. En principe, elles devraient compter sur des équipes d'experts

(de même qu'elles devraient pouvoir s'appuyer sur les connaissances et l'expérience d'autres services réglementaires et de surveillance, avec lesquels elles partageraient leur expérience). Dans la pratique, cependant, peu ont développé des telles capacités internes de recherche et de production de connaissances.

Ces différentes entraves aux capacités des AAC constituent également un défi au regard de leurs procédures de S&E et de leur évaluabilité. En fait, les AAC ont certains points communs sur lesquels les évaluateurs peuvent s'appuyer, mais ils doivent également impérativement savoir reconnaître les subtiles différences existant entre les mandats et les fonctions des unes et des autres, ainsi que les facteurs qui sont généralement causes d'échec institutionnel. Cela permettra d'améliorer l'évaluabilité des agences et de favoriser l'élaboration d'évaluations factuelles.

4.2 Améliorer l'évaluabilité des AAC

L'évaluabilité des AAC devrait être assurée par l'intégration systématique de certaines caractéristiques à leurs programmes et à leur schéma institutionnel. Lorsqu'une évaluation est amorcée, il est souvent trop tard pour remédier à des problèmes de fond. Plusieurs facteurs peuvent contribuer à améliorer l'évaluabilité des AAC, parmi lesquels une grande partie dépend de l'implication du personnel des AAC et des donateurs ainsi que des ressources que les agences peuvent consacrer au renforcement de leurs capacités internes en matière de S&E. Les AAC dont les capacités sont insuffisantes devraient entreprendre un certain nombre d'activités (de préférence avec le soutien des donateurs et de formateurs / d'animateurs expérimentés) avant le début de l'évaluation.

Dans un recueil sur la question de l'étude de l'évaluabilité, la Commission européenne (2009) recommande la mise en œuvre de cinq mesures favorisant l'évaluabilité :

- Enoncer clairement la logique d'intervention et la chaîne de résultats. Autrement dit, clarifier les objectifs stratégiques (c'est-à-dire la finalité), les mécanismes et les moyens prévus, les résultats attendus et les postulats sur lesquels se base l'intervention.
- Evaluer dans quelle mesure la conception, la stratégie, les ressources et les mécanismes de mise en œuvre du programme concordent avec la logique d'intervention.
- Evaluer dans quelle mesure le programme est susceptible d'atteindre les objectifs de la stratégie globale ou du plan d'action prioritaire pour lesquels il a été conçu.
- Evaluer la disponibilité des informations nécessaires, et notamment celle des sources de données primaires et secondaires, ainsi que le coût et la difficulté d'accès probables à ces renseignements, en incluant la disposition des relais à coopérer.
- Envisager dans quelle mesure une évaluation est susceptible d'améliorer les performances d'un programme et de contribuer à son succès et considérer si ces apports sont proportionnels aux coûts et aux efforts qui devront probablement être investis dans l'évaluation.

Les trois premières mesures visent à établir clairement les théories du changement et les indicateurs de performance des AAC. Nous fournirons des indications à ce sujet dans le chapitre 4.2.1. Les deux dernières étapes dépendent à la fois de l'existence d'un système de S&E fonctionnel et du degré de capitalisation des résultats au sein de l'AAC. Nous aborderons ce point dans le chapitre 4.2.2.

Il est à noter que ces exercices peuvent entrer dans le cadre de la mission de l'évaluateur, ce dernier pouvant en effet être chargé d'établir une théorie du changement et des indicateurs en amont de l'évaluation. Dans ce cas, les chapitres 4.2.1 et 4.2.2 relèveraient alors des chapitres 5 et 6 relatifs à l'amélioration de l'évaluation des résultats et des impacts. Il est cependant préférable que les AAC et les donateurs s'impliquent activement dans l'élaboration des théories du changement, des indicateurs et des systèmes de S&E, car ils maîtrisent mieux la logique d'intervention, et peuvent par la même

occasion renforcer leurs capacités institutionnelles. L'évaluateur doit intervenir en tant qu'animateur et formateur plutôt que comme responsable de ces étapes au premier chef.

4.2.1 Etablir des théories du changement et des indicateurs de performance précis

Dans ce chapitre, nous verrons pourquoi il est important que les AAC se dotent de buts et d'objectifs immédiats et comment il est possible de déterminer si une organisation suit la bonne voie pour atteindre ses objectifs. Bien que ces questions relèvent d'une approche plus globale de gestion axée sur les performances, nous nous concentrerons ici plus particulièrement sur les aspects relatifs au S&E. L'instauration de théories du changement et de bons indicateurs améliore considérablement l'évaluabilité d'une AAC et les chances d'obtenir des évaluations de qualité. Cette partie se rapporte donc également au chapitre 5 sur les méthodes d'amélioration des évaluations.

En premier lieu, nous présenterons les principaux défis auxquels sont confrontées les AAC qui souhaitent se fixer des modèles logiques et des théories du changement ainsi que des indicateurs de performance. Nous dresserons ensuite une liste de recommandations pour les aider dans leur tâche.

Les principaux défis auxquels sont confrontées les AAC dans la définition de leurs objectifs, leurs théories du changement et leurs indicateurs

Très peu d'AAC définissent leurs objectifs et les produits ou les réalisations attendus dans leur plan stratégique. Il est par conséquent très difficile d'évaluer leur réussite. Comme l'a souligné Meagher, « Il est rare qu'une nouvelle agence exprime – et veuille exprimer – clairement dans quelle mesure ses objectifs reflètent un désir de changement systémique plutôt qu'une manœuvre politique, ce qui pose un dilemme pour en définir et en évaluer le succès. Quel point de référence utiliser alors ? » (2002, 7).

Les AAC auraient tout intérêt à détailler leurs objectifs au maximum. Ainsi, plutôt que de viser la réduction de la corruption à l'échelon national, elles gagneraient à subdiviser cet objectif en éléments plus précis et mieux observables. Pour ce faire, elles peuvent notamment spécifier les *sphères de corruption*. Par exemple, visent-elles le secteur public ou privé ? Dans le cas du secteur public, ciblent-elles le gouvernement central ou décentralisé ? Si la cible est le gouvernement central, des ministères sont-ils visés en particulier ? Il conviendrait également qu'elles fassent une distinction entre les différents *types de corruption* ciblés : bureaucratique ou politique, petite ou grande, etc. De ces objectifs découleront les théories du changement et les indicateurs à mettre en place. Lorsque les objectifs sont trop généraux et constituent plus une démarche politique que des cibles réalistes, les évaluateurs sont alors confrontés à une tâche difficile, pour ne pas dire impossible.

Les AAC tendent à suivre des objectifs très généraux au niveau national, c'est pourquoi les différents groupes de bénéficiaires sont très difficiles à distinguer. Par ailleurs, contribuer à faire reculer la corruption, objectif classique des AAC, est par définition difficile à mesurer. Même lorsque le recul de la corruption peut se vérifier, il ne découle pas nécessairement du travail d'une AAC, d'autant plus que, dans un système national de lutte pour l'intégrité, les agences ne sont généralement pas les seuls intervenants. Comme l'explique Meagher (2002, 10),

Les antécédents positifs d'une AAC – c'est-à-dire, la qualité de ses méthodes de travail les preuves de son efficacité – ne permettent pas à eux seuls d'évaluer la valeur ajoutée nette de l'agence. Si cette AAC en question n'existait pas, une autre agence exercerait probablement ses fonctions de prévention, de dissuasion et de sensibilisation. Les AAC le font-elles mieux, au point d'être plus performantes que d'autres organismes dans leurs réalisations et leurs produits ? Mais alors, comment y parviennent-elles, et comment le savoir ?

La situation révèle ici un problème d'attribution, en plus de l'impossibilité de cerner le groupe ciblé, sans compter que dans certains cas, les objectifs ne sont en plus pas définis en bonne et due forme ou que, lorsqu'ils le sont, ils diffèrent parfois considérablement des impacts réels recherchés.

Les AAC utilisent souvent les produits réalisés pour faire valoir leurs résultats. Or ce sont des substituts inadaptés. En aucun cas la mesure d'un produit, comme par exemple le nombre d'affaires instruites, ne peut être considérée comme révélatrice d'un résultat. Elle n'indique que l'intensité des activités, et n'est pas la marque des réalisations ou des changements opérés. L'utilisation d'indicateurs quantitatifs de produits exige également « de considérer d'un point de vue qualitatif si l'agence et les activités sont bien ciblées et si les résultats ne pourraient pas être plus bénéfiques » (USAID, 2006, 24).

La survie des AAC, comme celle d'autres institutions publiques, tient à d'autres facteurs qu'à leur seul niveau de performance. Les AAC « sont créées et dissoutes par décision politique, et pas toujours en connaissance de cause » (De Sousa, Hindness et Larmour, 2009, 14). En outre, elles sont généralement sous contrôle rapproché du gouvernement et utilisées à des fins politiques (par exemple, pour attaquer l'opposition ou pour séduire le public et la communauté internationale). Leur évaluation n'en est que plus complexe. Pour compenser – au moins en partie – la dimension politique des AAC, les évaluations peuvent établir un lien entre les performances attendues et la disponibilité des ressources. On ne peut attendre de grands résultats de la part d'agences qui disposent de personnel insuffisant et/ou peu qualifié, manquent de ressources et ne sont pas investies des compétences juridiques nécessaires pour engager des poursuites en cas de délit.

Ces différents obstacles rendent les évaluations difficiles, c'est pourquoi ces dernières doivent être fondées sur des objectifs stratégiques clairs ainsi que des buts, des finalités et des produits bien définis. Entreprendre une évaluation alors que les buts, les objectifs et les indicateurs de l'AAC sont trop larges, trop généraux, trop vagues ou tout simplement inexistant relèvent de la gageure. C'est également le cas lorsque l'objectif affiché diffère de l'objectif effectif déterminé par les réalités politiques. Pour garantir des évaluations productives, les dirigeants d'AAC doivent avant tout identifier la vision et la mission réelles de l'organisation ainsi que les objectifs et les indicateurs qui en découlent.

Les grandes lignes du modèle logique et de la théorie du changement pour une AAC

Avant d'entreprendre une évaluation, une AAC devrait s'assurer qu'elle compte bien sur une théorie du changement raisonnée, cohérente et clairement articulée sur laquelle baser ses interventions. Pour la représenter, elle peut utiliser un modèle logique ou une chaîne de résultats. C'est l'une des composantes fondamentales de notre méthodologie d'évaluation des impacts (voir chapitre 6). Pour reprendre la terminologie définie plus haut, on peut dire que les organisations mettent en œuvre des programmes pour obtenir un ou plusieurs *impacts*. Ces programmes reposent sur des *moyens* et des *activités*, qui aboutissent à des *produits*, lesquels doivent mener aux *réalisations* escomptés. Sur le long terme, ces derniers se reflèteront dans les *impacts*.

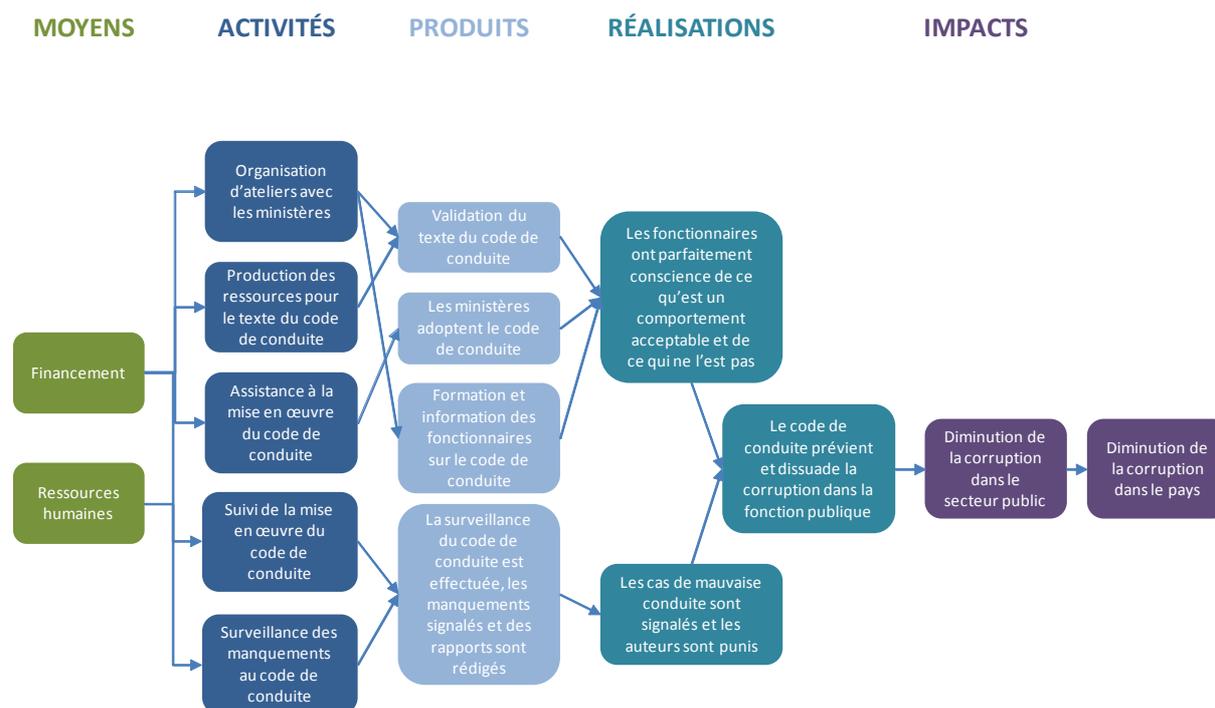
Comme le montre le schéma 2, des liens logiques doivent apparaître entre les moyens, les activités, les produits, les réalisations et les impacts. Des produits ne contribuant à aucun résultat sont probablement superflus ou révélateurs d'un modèle mal construit. Le même ordre sous-tend les cadres logiques : ceux qui sont bien conçus voient les produits mener à des réalisations (effets), lesquels occasionnent des impacts (buts).

Lorsqu'il n'existe pas de modèle logique préalable, l'évaluateur peut en élaborer un de substitution et obtenir l'autorisation de procéder à l'évaluation sur cette base. Ce n'est toutefois pas une solution recommandée.

Les AAC étant souvent des organisations polyvalentes qui poursuivent des objectifs multiples, il peut être nécessaire d'élaborer plusieurs modèles logiques pour une même agence. Généralement, si cette dernière mène des activités aux résultats et/ou aux impacts très différents, il est en effet sans doute préférable d'établir des modèles logiques séparés, puisque l'AAC base ses interventions sur des théories du changement différentes.

Pour illustrer ce point, nous allons voir un exemple fictif de modèle logique dans le cas d'une AAC chargée de mettre en œuvre un code de conduite dans le secteur public.

Schéma 2. Théorie du changement pour la mise en œuvre d'un code de conduite

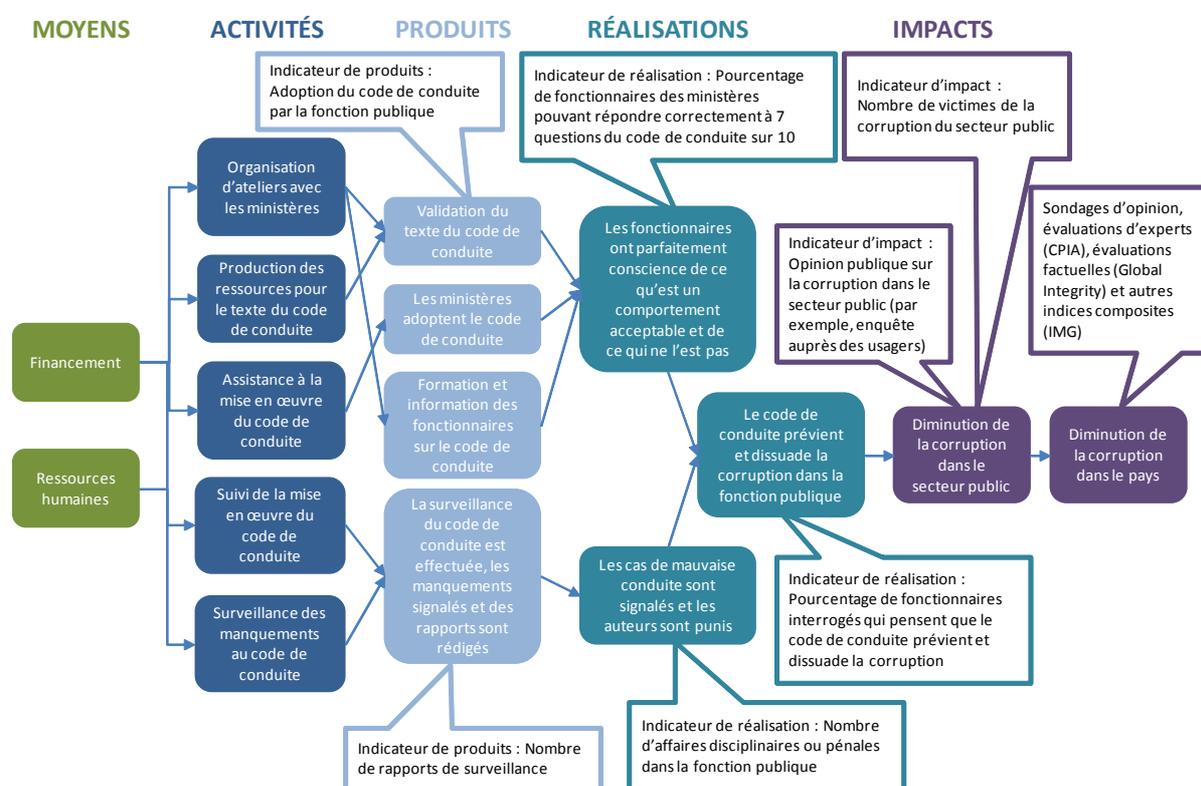


Ce modèle logique démontre comment le travail effectué autour d'un code de conduite peut aboutir à l'impact désiré, à savoir, la diminution de la corruption. En spécifiant le double effet préventif et répressif que le code de conduite est supposé avoir, il clarifie la théorie du changement. Certains postulats à l'origine de cette théorie du changement sont toutefois sujets à question. Par exemple, rien ne dit que les fonctionnaires conscients des normes de bonne conduite et des sanctions appliquées en cas d'infraction soient pour autant réellement dissuadés de se livrer à des actes de corruption. Tout dépend du contexte. Toutefois, le fait d'aborder le sujet dans une théorie du changement explicite permet au moins de soulever la question. Il n'existe pas de modèle standard de théorie du changement mais, en principe, cette dernière devrait refléter au mieux les méthodes d'intervention, être élaborée de manière participative et être acceptée par toutes les parties prenantes.

En phase de planification d'un programme, l'élaboration d'une théorie du changement constitue un exercice hypothétique qui vise à anticiper le meilleur moyen d'obtenir les effets escomptés. Lorsqu'une théorie du changement est utilisée à des fins d'évaluation, elle doit être aussi proche de la réalité que possible, afin de permettre une comparaison claire avec la théorie hypothétique précédente, établie en amont du programme. Les théories du changement ne sont généralement pas formulées de manière explicite, mais l'on sait que toute intervention réfléchie se base toujours sur une certaine logique.

Une fois le modèle logique construit et la théorie du changement élaborée, il convient de procéder au choix des paramètres clés à évaluer. Des indicateurs peuvent être créés pour chacun d'entre eux, qui seront facilement localisables sur le modèle logique, que ce soit au niveau des produits, des réalisations ou des impacts. Dans le schéma 3, nous allons voir des exemples d'indicateurs correspondants aux différents niveaux du modèle logique. Notons que jusqu'au niveau des résultats inclus, les indicateurs s'appliquent au projet (ici, la mise en œuvre d'un code de conduite), mais qu'au niveau des impacts, ils sont plus généraux et moins spécifiques à l'intervention. L'évaluation d'une AAC ne peut donc se baser uniquement sur les indicateurs d'impact, car ceux-ci ne permettent pas de savoir si un impact observé est dû à l'efficacité de l'intervention de l'agence ou plutôt à un événement extérieur ou à une autre organisation. Dans le schéma 3, par exemple, l'un des indicateurs d'impact se rapporte à l'évolution de l'opinion des usagers des services publics sur le niveau de corruption dans le secteur public. Cet indicateur peut devenir positif du fait de l'instauration d'un nouveau service d'assistance téléphonique du gouvernement pour signaler d'éventuelles infractions, et non pas grâce au code de conduite. C'est seulement en intégrant les indicateurs de résultats à l'évaluation que l'on peut comprendre si le code de conduite a contribué à réduire le niveau général de corruption.

Schéma 3. Théorie du changement avec indicateurs



La nécessité de disposer d'indicateurs spécifiques à l'intervention requiert en outre que ces derniers soient élaborés au niveau du projet ou de l'organisation, tout comme la collecte des données. Des indicateurs universels ou mondiaux tels que l'IPC, les Évaluations de la politique et des institutions nationales (CPIA) de la Banque mondiale ou les Indicateurs mondiaux de gouvernance (IMG) de l'Institut de la Banque mondiale ne sont pas suffisants.

Etablir des indicateurs et des références

Pour identifier des indicateurs à la fois pertinents et mesurables, les AAC devraient avant toute chose *entreprendre une consultation des parties prenantes* impliquant les donateurs, leur propre personnel, les organismes gouvernementaux, les experts et la société civile. Cette démarche favorise une compréhension commune et facilite l'adhésion aux indicateurs de performance ; elle améliore la

qualité des indicateurs et permet de garantir qu'à la fin de la procédure, toutes les parties en jeu souscriront aux conclusions de l'évaluation.

En second lieu, il convient de *veiller à ce que les indicateurs soient mesurables*. A cet effet, il est utile de procéder à un examen des sources de données disponibles et d'en évaluer la pertinence, l'applicabilité, la validité et la fiabilité. Lorsque c'est possible, les évaluateurs doivent éviter de repartir de zéro et utiliser les sources de données existantes. Les AAC peuvent cependant contribuer à la tâche en rassemblant ces données dans une base de données unique (nous détaillerons plus loin le management électronique du système d'information). De manière générale, le personnel des AAC chargé du S&E devrait travailler en collaboration avec le bureau national des statistiques (pour les statistiques générales), le pouvoir judiciaire (pour les statistiques sur la criminalité) et les ONG de lutte contre la corruption (pour les enquêtes sur la perception de la corruption).

L'élaboration d'indicateurs est une étape cruciale dans le développement d'un bon système de S&E axé sur les performances, car elle détermine la collecte, l'analyse et la transmission des données qui s'ensuivent. Il est donc conseillé de consulter avec soin les recommandations sur ce sujet qui, même si elles ne sont pas spécifiques au domaine de la lutte contre la corruption, sont universellement applicables. En anglais, il existe un moyen mnémotechnique bien connu pour la préparation d'indicateurs : SMART, selon lequel les indicateurs doivent être spécifiques (**S**pecific), mesurables (**M**asurable), atteignables (**A**ttainable), pertinents (**R**elevant) et limités dans le temps (**T**ime-bound). Schiavo-Campo (1999) a quant à lui conçu un autre ensemble de critères à partir de l'acronyme CREAM, renvoyant à des indicateurs à la fois clairs, précis et non équivoques (**C**lear), pertinents et en rapport avec le sujet (**R**elevant), économiques et permettant de produire des données à un coût raisonnable (**E**conomic), adéquats et constituant une base suffisante pour l'évaluation des performances (**A**dequate) et enfin, contrôlables et susceptibles d'être validés de manière indépendante (**M**onitorable). Chacune de ces deux méthodes est utile et doit être davantage considérée comme un ensemble idéal de critères que comme un modèle exclusif d'indicateurs.

Le choix des indicateurs dépend des systèmes de données, des méthodes de collecte des informations, des ressources et des capacités des AAC. Les colonnes « sources de vérifications » et « commentaires, postulats » du catalogue d'indicateurs (annexe 1) doivent autant que possible faire apparaître tout éventuel système, méthodologie ou capacité spéciale, ou tout autre besoin particulier, qu'un indicateur est susceptible de requérir. Enfin, il est fortement souhaitable que les indicateurs soient adaptés et affinés par les mêmes personnes qui les utiliseront, aussi bien avant le début de la collecte des données qu'au titre d'un exercice continu. Comme le suggèrent Morra Imas et Rist (2009, 118), cette tâche peut être entreprise au cours d'un atelier, au cours duquel les facteurs suivants peuvent être détaillés pour chaque indicateur :

1. Les sources des données
2. La méthode de collecte des données
3. Les responsables de la collecte des données
4. La fréquence de collecte des données
5. Le coût de la collecte des données
6. Les difficultés à collecter les données
7. Les responsables de l'analyse et de la transmission des données
8. Le type d'analyse (comprenant les limites et les avantages des approches alternatives)
9. Les destinataires des données

Aucune des évaluations passées en revue dans le cadre de notre inventaire ne prévoit d'indicateurs satisfaisants. La conception d'indicateurs est fondamentale pour l'intégralité de la procédure

d'évaluation, c'est pourquoi nous consacrerons le chapitre 7 et l'annexe 1 à ce sujet, que nous illustrerons par différents exemples et accompagnerons de recommandations. Une fois les indicateurs identifiés et formulés, l'étape suivante est l'établissement de données de référence (baseline data), de repères (benchmarks) et d'exemples de scénarios contrefactuels.

Il est important que les donateurs contribuent à définir les indicateurs sur lesquels ils baseront leurs évaluations, plutôt que de les imposer de l'extérieur aux AAC. Ces dernières devraient chacune disposer d'un système de S&E, et non pas se voir appliquer un système différent par chaque bailleur. Le mieux est de sélectionner quelques indicateurs clés seulement pour mesurer leurs performances. Le choix des indicateurs peut être effectué en fonction des motifs de préoccupation de la population (selon les secteurs, régions ou types de corruption concernés) ou du niveau auquel l'Etat est le plus susceptible de perdre des revenus (au niveau du gouvernement central ou local, de secteurs spécifiques ou d'institutions particulières).

Pour être en mesure d'établir des données de référence pour les indicateurs définis, les AAC devraient réaliser ou mandater des enquêtes, des évaluations qualitatives et des évaluations de risques, par exemple sur l'opinion du public sur leur travail ou sur le nombre annuel d'enquêtes menées sur des affaires de corruption. Une AAC qui souhaite améliorer l'une de ses procédures peut par exemple choisir le délai de traitement des dossiers comme indicateur, dresser une représentation graphique du temps de traitement de différents types de dossiers et assurer un suivi régulier de l'amélioration de ses performances. Les données de référence peuvent également permettre une comparaison des performances de différentes AAC dans l'accomplissement de leurs fonctions de base. La diversité des agences (que nous avons évoquée dans le chapitre 4.1.2) est toutefois telle qu'il est très difficile, sur le plan méthodologique, d'effectuer des comparaisons à l'échelle internationale.

4.2.2 La création et la gestion de systèmes de S&E fonctionnels

Les évaluations dépendent d'une multitude de sources de données crédibles. Actuellement, la plupart des évaluations d'AAC se basent exclusivement sur des entretiens avec des personnes ressources, ce qui limite considérablement (a) la formulation de conclusions sérieuses et tangibles sur leurs performances et leurs impacts, (b) la production de résultats susceptibles d'être généralisés et utilisés à des fins de comparaison, et (c) la création de données factuelles à soumettre aux responsables politiques. Un service de S&E opérationnel et bien géré constitue souvent une mine d'or pour les évaluateurs, notamment si les indicateurs nécessaires sont définis dès le départ et que le service en question procède à un recueil régulier des données et complète ainsi ses propres rapports avec, par exemple, les données issues de petites enquêtes.

La différence entre les produits et les réalisations

Le propre d'un système de S&E axé sur les performances est sa capacité à dépasser le niveau des produits (outputs) pour parvenir à celui des réalisations (outcomes). Les évaluateurs et le personnel des organismes donateurs et des AAC ont malheureusement tendance à considérer que les produits peuvent se substituer aux réalisations. Les deux concepts sont pourtant radicalement différents, c'est pourquoi il convient de bien les dissocier. Les AAC utilisent les moyens qu'elles ont à leur disposition pour entreprendre des activités grâce auxquelles elles réalisent des produits, lesquels sont supposés générer les réalisations escomptées. Sur le long terme, les réalisations devraient se refléter dans des performances ou des impacts plus généraux. Les produits relèvent entièrement des choix des AAC, alors que les réalisations sont indépendantes de leur volonté. Autrement dit, les produits peuvent entraîner des réalisations aussi bien prévisibles qu'inattendues. Cependant, si le modèle logique ou la théorie du changement est valable, les indicateurs de réalisations sont censés varier. Le volume de travail accompli ou les résultantes immédiates des activités entreprises sont par exemple des produits. Ce sont des produits et des services fournis par les AAC et dont la variation est reflétée dans les indicateurs de produits, comme par exemple le nombre de formations effectuées. Les réalisations, en revanche, sont le fruit de l'usage ou de l'application des produits, comme par exemple, des

fonctionnaires mettant en application les connaissances acquises au cours de formations anti-corruption. Enfin, les réalisations révèlent les bénéfices qu'un programme donné aura sur le groupe cible et traduisent notamment l'évolution de la situation, des habitudes, des perceptions, etc.

Les composantes d'un système de S&E

Avant de concevoir un cadre de S&E, il est indispensable de s'assurer que les systèmes soient en place et les responsabilités assignées pour la collecte, le suivi, l'analyse et le stockage des données nécessaires à l'élaboration des indicateurs identifiés. Il est souvent nécessaire d'affecter une personne aux activités de S&E à plein temps. Cela permet en effet d'assurer un suivi continu, de limiter la dépendance des agences vis-à-vis de consultants à court terme plus ou moins compétents et d'alimenter la mémoire institutionnelle. La personne en question devra remplir des fonctions de collecte de données et de validation de « conformité », mais également collaborer avec la direction afin de veiller à ce que les plans, les activités et les stratégies mis en place contribuent bien aux fins et aux objectifs de l'AAC.

Le système de S&E doit stipuler les besoins en informations, les méthodes et les intervalles de collecte des données et les cadres d'analyse. De même, il doit fournir un système de gestion de l'information pour le stockage, la collation, l'analyse, l'utilisation et la diffusion des données. Le cas échéant, le système de S&E doit intégrer le cadre logique convenu avec les donateurs, lequel doit être pratique et simple à utiliser, mais capable de saisir tous les indicateurs de performance nécessaires. Le système de S&E doit par ailleurs clairement définir les responsabilités de chacun des services de l'AAC dans les différentes activités de S&E.

Pour garantir une collecte systématique des données, le système de S&E devrait absolument prévoir le management électronique des systèmes d'information (MSI) permettant de stocker les données collectées lors des opérations de suivi. Tout système de S&E bien conçu doit permettre la collecte de volumes importants de données et l'exploitation de grandes quantités d'informations. Pour pouvoir traiter ces informations en temps réel, mais aussi les utiliser à plus long terme, les AAC doivent appliquer des procédures de gestion rigoureuses et intégrer leurs données à un système solide de stockage et de recherche. Elles peuvent pour cela élaborer une simple base de données sur Excel ou sur d'autres programmes similaires. Celles qui traitent un grand nombre de données spatiales préféreront peut-être, si elles en ont les moyens, utiliser des systèmes plus élaborés permettant d'effectuer des renvois et d'exploiter les données de systèmes d'information géographique. Avant tout, le MSI doit permettre la saisie et la recherche de données dans une optique de partage, de remontée, d'utilisation et de diffusion des informations.

Le personnel des AAC doit comprendre que l'honnêteté est le pilier des systèmes de S&E, et doit éviter de dissimuler ou de supprimer d'éventuelles erreurs ou défaillances. Les dirigeants d'agences doivent par ailleurs s'attacher à créer une culture de l'apprentissage et à encourager une communication honnête des résultats, qu'ils soient positifs ou négatifs. Les AAC ne pourront tirer de leçons de leurs activités que si elles rendent systématiquement compte de leur travail et qu'elles fondent leurs décisions sur les conclusions des analyses effectuées de leurs rapports.

L'élaboration d'un système de S&E au sein d'une AAC doit respecter certains principes de base :

- Définir les priorités. Cibler quelques indicateurs au niveau des résultats et des impacts, surtout si l'AAC doit collecter des données externes.
- Développer le sens critique et les capacités d'analyse des personnes impliquées, y compris des dirigeants et des administrateurs de programme ainsi que du responsable du S&E.
- Mettre l'accent sur l'analyse et les prises de décisions.
- S'assurer de traiter les variations négatives et imprévues en plus des évolutions prévues ou planifiées.

- Structurer le système autour de l'apprentissage/capitalisation de l'expérience.
- Produire des informations utiles à toutes les catégories de parties prenantes.
- Rendre des rapports exhaustifs réguliers, basés sur un suivi constant des indicateurs sélectionnés.

L'élaboration, la mise en œuvre et la viabilité d'un système de S&E dépendent des composantes suivantes, basées sur les principes susmentionnés :

- *Une vision* : comprendre comment les informations générées par le S&E peuvent contribuer à la gestion des AAC et alimenter les processus décisionnels. Les AAC doivent pour cela développer un leadership stratégique et acquérir une bonne compréhension des concepts de base du système et de l'utilisation qu'elles peuvent en faire.
- *Un cadre favorable* : s'engager non seulement à lancer un système de S&E, mais également à le maintenir sur le long terme.
- *Les capacités techniques et l'infrastructure pour traiter les informations du S&E* : cela comprend des systèmes crédibles et adéquats de collecte d'informations et de données (comme, par exemple, de petites enquêtes), ainsi que du personnel qualifié capable de collecter, d'analyser et de transmettre les progrès effectués par rapport aux objectifs visés. La mise en place d'une base de données peut demander la formation des spécialistes en S&E et du personnel technique afin d'assurer un bon usage du système. C'est un processus long et répétitif, mais l'apprentissage continu est vecteur d'amélioration du système de S&E.
- *L'infrastructure pour demander et utiliser les informations du S&E* : cela requiert des mécanismes à la fois informels (par exemple, des espaces de réflexion) et formels (par exemple, un MSI intégré) de transmission et de partage de l'information générée par le S&E.

En fait, les systèmes de S&E des AAC doivent rester simples. Ils peuvent se baser sur des processus clairs et innovants et reposer sur une remontée de l'information participative avec les parties prenantes, sur une analyse approfondie des données et sur un ensemble d'instruments favorisant l'engagement actif des acteurs en jeu. Les ressources et le temps consacrés au développement d'un tel système constitueraient un bon investissement.

La gestion des procédures d'évaluation

Les employés des services de S&E des AAC devraient être chargés de gérer les procédures d'évaluation, si possible en collaboration avec leurs homologues des organismes donateurs. Plusieurs conditions peuvent contribuer à la réalisation d'évaluations productives. Tout d'abord, les AAC devraient dès le début définir clairement leurs attentes en matière de critères de réussite et déterminer si elles disposent des ressources et du mandat nécessaires à la satisfaction de ces critères. Elles doivent par ailleurs s'assurer de pouvoir établir une logique de causalité ou une théorie du changement. En second lieu, elles doivent concentrer leurs efforts sur l'évaluation des réalisations plutôt que

ENCADRE 4. LES ELEMENTS CONSTITUTIFS D'UN BON SYSTEME DE S&E

- Une description claire d'*objectifs mesurables*
- Un ensemble structuré d'*indicateurs* clairement liés aux objectifs mesurable
- Des dispositions pour collecter et gérer les *rapports de projets* de manière à ce que les données requises pour les indicateurs soient disponibles à moindre coût
- Des dispositions pour *collecter, analyser et transmettre les données*
- Des dispositions permettant de veiller à ce que les résultats et les recommandations *influencent bien les processus décisionnels*

sur la mesure des produits et des activités, qui sont de piètres révélateurs d'impact. Troisièmement, elles doivent veiller à ne pas surcharger les évaluations en indicateurs. Quatrièmement, elles doivent intégrer le fait que les évaluations gagnent à reposer sur des méthodologies mixtes, et devraient donc utiliser des indicateurs aussi bien quantitatifs que qualitatifs. En dernier lieu, les évaluations devraient être mises à la disposition du public. Les commentaires et opinions des citoyens, des donateurs et de toute autre partie prenante sont en effet indispensables à l'amélioration des indicateurs choisis, de la cible de l'analyse, de l'étude de scénarios contrefactuels et enfin, de la qualité du travail des AAC.

Promouvoir la capitalisation de l'expérience

Un système de S&E ne doit pas uniquement répondre à des impératifs de redevabilité, mais doit également viser la capitalisation de l'expérience. Une démarche de S&E qui se limite à rendre des comptes relève de l'audit, se résume souvent à l'emploi de listes de contrôle et consiste à vérifier que l'AAC ait dépensé ses ressources (moyens) dans les activités prévues et réalisé les produits attendus (produits). Une AAC qui cherche à capitaliser son expérience privilégie les réalisations et les impacts. Elle utilise les données collectées pour tenter d'expliquer les bonnes ou les mauvaises performances réalisées et en déduire des propositions d'amélioration. Dans le chapitre 8, nous reviendrons sur l'importance de la capitalisation des connaissances tirées du S&E et fournirons un certain nombre de recommandations aux AAC qui souhaitent adopter une approche systématisée de la capitalisation dans leurs cadres de S&E.

5. Comment réaliser de meilleures évaluations ?

Dans ce chapitre, nous présenterons les principes de meilleures pratiques en matière d'évaluation. Ces principes, déjà utilisés dans d'autres secteurs depuis des dizaines d'années, devraient permettre au personnel des AAC, aux donateurs et aux évaluateurs de voir évoluer leurs pratiques et de réaliser des évaluations autrement qu'en « cochant des cases ». L'objectif est de revaloriser les concepts de « réalisations » et d'« impacts » et d'effectuer les investissements nécessaires en travail et en ressources pour effectuer une évaluation approfondie de ces derniers.

Comme nous l'avons vu dans le chapitre 2, les évaluations d'AAC ne tirent généralement aucune conclusions factuelles sur les résultats et les impacts des agences, car elles tendent à cibler uniquement les performances organisationnelles que ces dernières réalisent au niveau des moyens, des activités et des produits. Une AAC qui semble avoir de bons produits peut donc en réalité avoir des impacts insignifiants et n'induire aucun changement positif durable. Nous allons voir ici comment effectuer une meilleure évaluation des réalisations et des impacts des AAC afin que ces dernières soient en mesure d'accomplir un travail de meilleure qualité et de satisfaire aux pressions extérieures en matière de résultats.

Lorsque les évaluations sont d'approche plus complexe, comme c'est le cas dans les domaines de la lutte contre la corruption, des droits de l'homme, de l'autonomisation des femmes, etc., il convient avant tout de revenir sur le problème que l'intervention, en l'occurrence, la création d'une AAC, était censée résoudre. Cette position négative servira alors de référent pour mesurer les variations enregistrées. Cela implique de reconnaître que les méthodologies d'évaluation « superficielles » ne sont pas suffisantes et que des stratégies d'évaluation complexes doivent être mises en place.¹⁵

Ce chapitre s'inspire largement des critères d'évaluation de l'OCDE, lesquels sont ici toutefois adaptés aux AAC. Les évaluateurs expérimentés y trouveront certainement des répétitions. Si certaines informations peuvent leur sembler évidentes, nous avons choisi de les reproduire afin de nous assurer que toutes les parties prenantes parlent le même langage et adhèrent aux mêmes critères.

Ce chapitre vise avant tout les évaluateurs et le personnel des AAC. Les donateurs peuvent cependant tirer profit des directives sur la conception et la réalisation de bonnes évaluations, lesquelles reposent notamment sur la définition de termes de référence pour chaque évaluation et sur l'obtention d'une garantie de qualité.

La qualité de l'évaluation d'une AAC repose sur les sept principaux critères suivants :

- La finalité et la justification de l'évaluation
- La composition de l'équipe d'évaluation
- Le champ couvert, les principes généraux, la méthodologie et les questions de l'évaluation
- Les méthodes d'analyse employées pour l'évaluation
- La participation et la transparence
- La garantie de qualité
- Les aspects budgétaires et financiers du S&E

¹⁵ Notre définition d'une évaluation complexe se base sur celle d'EuropeAid, à savoir, une évaluation qui implique de multiples activités, des groupes cibles hétérogènes, de nombreux effets escomptés et différents impacts attendus (Commission européenne, 2006b, 17-18).

5.1 La finalité et la justification de l'évaluation

Pour toute évaluation, la première étape importante est de s'accorder sur l'objet et la finalité de l'exercice, qui doivent être transmis à toutes les parties prenantes et à tout le personnel de l'ACC, et figurer dans les termes de référence de l'évaluation. Ces derniers doivent également faire apparaître les objectifs de l'évaluation, les modalités de traitement des résultats obtenus, le degré de précision requis et enfin, le type de décisions stratégiques et/ou de gestion que les conclusions doivent permettre d'alimenter.

Par ailleurs, il est également important de décrire le contexte de l'évaluation et d'indiquer qui en est l'initiateur. Actuellement, les évaluations des performances des AAC sont généralement engagées soit par l'agence elle-même, soit par le gouvernement, soit par les organismes donateurs.

Dans un deuxième temps, les objectifs de l'évaluation doivent être traduits en questions pertinentes et spécifiques à l'évaluation (voir chapitre 5.3).

ENCADRE 5. DEFINIR LA JUSTIFICATION, LA FINALITE ET LES OBJECTIFS D'UNE EVALUATION

D'après les Normes de qualité pour l'évaluation du développement du CAD/OCDE :

La justification, la finalité et l'utilisation souhaitée de l'évaluation [devraient être] clairement formulées en répondant aux questions suivantes : pourquoi l'évaluation est-elle entreprise à ce moment précis ? Pourquoi et pour qui est-elle entreprise ? De quelle manière l'évaluation sera-t-elle utilisée pour en retirer des leçons et pour rendre des comptes ?

Par exemple, la finalité générale de l'évaluation peut être de :

- contribuer à améliorer une politique, une procédure ou une technique de développement,
- décider de la poursuite ou de l'abandon d'un projet ou d'un programme,
- rendre compte de l'utilisation des dépenses d'aide et des résultats obtenus aux parties prenantes et aux contribuables.

Les objectifs spécifiques de l'évaluation clarifient les enjeux sous-jacents de l'évaluation. Par exemple, pour :

- vérifier des résultats (produits, effets, impact) et apprécier l'efficacité, l'efficience, la pertinence et la durabilité d'une intervention particulière en matière de développement, [ou]
- dégager des constatations, des conclusions et des recommandations à propos d'une intervention particulière en matière de développement afin d'en tirer des enseignements pour de futures conceptions et mises en œuvre.

Source: CAD/OCDE, 2010b.

5.2 La composition de l'équipe d'évaluation

La complexité des évaluations d'AAC peut exiger l'intervention d'experts en techniques d'évaluation et de spécialistes du domaine dans lequel intervient l'agence, comme par exemple les déclarations de patrimoine, les conflits d'intérêts, etc. Les évaluateurs doivent être choisis sur des critères de crédibilité, d'expertise et d'indépendance vis-à-vis de l'organisation évaluée. Les membres de l'AAC

doivent être impliqués dans l'exercice, ce qui leur permettra de mieux appréhender les méthodes institutionnelles et d'adhérer aux résultats de l'évaluation.

Lorsqu'une évaluation est commanditée par des donateurs, il est plus délicat de constituer une équipe équilibrée d'experts externes et internes, car les donateurs « parachutent » souvent des experts internationaux. L'expérience nous permet d'affirmer que ce genre d'évaluations rapides, qui reposent souvent sur l'expertise d'un ou deux individus, produisent rarement des informations utiles. Il serait plus judicieux que les donateurs aident les AAC à améliorer leurs propres systèmes de S&E et à développer leurs capacités à s'auto-évaluer ou encore, qu'ils envisagent la possibilité de mandater des experts locaux ou régionaux pour réaliser l'évaluation sans en compromettre la qualité.

Il est important que les évaluateurs soient indépendants de l'organisation évaluée, de ses fonctions stratégiques, opérationnelles et de gestion, ainsi que de ses bénéficiaires. Les éventuels conflits d'intérêts doivent être considérés de manière ouverte et honnête. Il est alarmant de constater que la majorité des évaluations passées en revue dans le chapitre 2 négligent complètement ces questions.

5.3 Le cadre, les principes généraux, la méthodologie et les questions de l'évaluation

5.3.1 Le champ couvert par l'évaluation

Les ambitions et les domaines d'intervention de l'évaluation, ainsi que les ressources qui y sont allouées, doivent être déterminés dès le début. Pour ce faire, l'AAC doit être en mesure de définir clairement sa vision et sa mission et de décrire sa théorie du changement, point que nous avons vu en détail dans le chapitre 4.2.1. Le cadre de l'évaluation définit la durée, les dépenses, la zone géographique, les groupes cibles, le schéma organisationnel, les modalités de mise en œuvre, le contexte politique et institutionnel et tous les autres aspects que l'évaluation doit couvrir (OCDE, 2010a, 11).

Il convient également de préciser si l'évaluateur doit effectuer une évaluation rapide ou une évaluation des capacités de l'organisation ou encore, une évaluation. Si l'exercice s'appuie sur les critères de pertinence, d'efficacité, d'efficience, d'impact et de durabilité de l'OCDE (voir plus loin), ces derniers seront-ils tous couverts ? Plus le cadre de l'évaluation sera étendu, plus il nécessitera de ressources.

5.3.2 Les principes et la méthodologie de l'évaluation

Pour exercer leur fonction, les évaluateurs doivent s'appuyer sur des principes et des critères qui favorisent la redevabilité, facilitent la comparabilité et améliorent la qualité et la fiabilité de leurs produits (Picciotto 2005, 30–59). La méthodologie de l'évaluation doit au minimum clarifier les points suivants :

- Le choix des domaines d'analyse
- La formulation des questions d'évaluation
- Le choix des méthodes

ENCADRE 6. LE PROJET DE RECHERCHE

Un projet de recherche fournit un cadre pour la collecte et l'analyse des données. Le choix du projet de recherche reflète les priorités accordées aux différents volets de la recherche, et notamment l'importance attribuée à (a) l'expression de liens de causalité entre les variables ; (b) la généralisation à des groupes d'individus plus larges que ceux qui entrent dans le cadre de l'étude ; (c) la compréhension des comportements et du sens de ces comportements dans des contextes spécifiques ; et (d) la vision temporelle (c'est-à-dire dans le temps) des phénomènes sociaux et de leurs liens.

Source: Bryman, 2004, 27.

- La formulation du projet de recherche et des techniques de collecte des données
- L'exécution de la collecte des données
- L'analyse des données
- L'interprétation des données
- Les conclusions¹⁶

La documentation en matière de méthode d'évaluation dans le domaine de l'aide au développement est abondante et détaillée. En présenter ici un tableau complet dépasserait le cadre de ce rapport. Cependant, l'examen que nous avons effectué des évaluations d'AAC a révélé qu'aucun des travaux réalisés n'adhérait intégralement aux principes directeurs établis. Pour cette raison, nous présenterons maintenant brièvement les principes d'une bonne évaluation, en nous basant sur les Critères d'évaluation développés en 1991 par le CAD/OCDE en matière d'assistance au développement (CAD/OCDE, 1986, 1991, 2000).

L'OCDE recommande de respecter les critères d'évaluation suivants : pertinence, efficacité, efficacité, impact et durabilité, qui constituent le pivot des termes de référence des évaluations. Les évaluations d'AAC ont tendance à cibler uniquement l'efficacité et, dans une certaine mesure, l'efficacité. Certaines, lorsqu'elles évaluent le mandat et les fonctions de l'AAC, abordent la question de la pertinence. Aucune ne remplit toutefois l'intégralité des critères de l'OCDE décrits dans l'encadré 7.

Les critères de l'OCDE n'ont pas à entrer tous les cinq dans le cadre de l'évaluation. Cela demanderait des ressources bien trop importantes et dépasserait les capacités de l'évaluateur, qui finirait par ne pouvoir en couvrir aucun. On considère généralement l'efficacité et l'impact comme les plus importants, mais tout dépend des problèmes spécifiques qui doivent être évalués. (Pour plus d'informations sur les évaluations d'impact, consulter le chapitre 6). Si le problème tient au fait qu'une AAC gaspille ses ressources par rapport aux autres AAC, l'étude devra alors avant tout porter sur le critère d'efficacité. Si c'est plutôt la pertinence de l'AAC vis-à-vis du cadre institutionnel général du pays qui est remise en question, l'évaluateur devra dans ce cas se concentrer sur le critère de la pertinence. Mesurer l'efficacité et l'impact requiert généralement l'usage de séries chronologiques de données, c'est pourquoi les premières évaluations devront sans doute plus se concentrer sur l'efficacité en attendant de disposer de données supplémentaires.

Les questions de l'évaluation doivent être formulées relativement tôt dans le processus et contribuer à la mise en place de la méthodologie. D'après l'OCDE,

La méthodologie comprend la spécification et la justification de la conception de l'évaluation et les techniques de collecte et d'analyse des données. La méthodologie choisie permet de répondre aux questions évaluatives en s'appuyant sur des preuves crédibles. Une distinction très claire est effectuée entre les différents niveaux de résultats (logique d'intervention comprenant une chaîne des résultats indiquant les moyens, les produits, les résultats et les impacts). (OCDE, 2010a, 24)

Comme nous le verrons dans le chapitre 5.4, l'élaboration d'une méthodologie exige de choisir des méthodes d'analyse.

¹⁶Bryman, 2004, 21–22.

ENCADRE 7. LES CRITERES DU CAD/OCDE EN MATIERE D'EVALUATION DE L'ASSISTANCE AU DEVELOPPEMENT

La pertinence : Mesure selon laquelle l'action de développement correspond aux priorités et aux stratégies du groupe cible, des bénéficiaires et des donateurs. Pour évaluer la pertinence d'un programme ou d'un projet, il est utile de réfléchir aux questions suivantes :

- Dans quelle mesure les objectifs du programme sont-ils encore appropriés ?
- Les activités et les produits du programme sont-ils cohérents avec les priorités globales et avec la réalisation de ses objectifs ?
- Les activités et les produits du programme sont-ils cohérents avec les impacts et les résultats souhaités ?

L'efficacité : Mesure selon laquelle les objectifs de l'action de développement ont été atteints. Pour évaluer l'efficacité d'un programme ou d'un projet, il est utile de réfléchir aux questions suivantes :

- Dans quelle mesure les objectifs visés (buts et réalisations) ont-ils été atteints [ou] sont-ils en passe d'être atteints ?
- Quels sont les principaux facteurs qui ont favorisé ou empêché la réussite des objectifs ?

L'efficience : Mesure les produits (outputs) – en termes qualitatifs et quantitatifs – par rapport aux moyens disponibles. Terme économique qui signifie que l'aide utilise les ressources les moins coûteuses possible pour atteindre les résultats souhaités. Cela requiert généralement de comparer des méthodes alternatives visant les mêmes produits afin de vérifier que le procédé le plus efficace a bien été employé. Pour évaluer l'efficience d'un programme ou d'un projet, il est utile de réfléchir aux questions suivantes :

- Le coût des activités a-t-il été optimisé ?
- Les produits ont-ils été accomplis à temps ?
- Le programme ou projet a-t-il été mis en œuvre de la façon la plus efficace possible par rapport à d'autres alternatives ?

L'impact : Effets à long terme, positifs et négatifs, induits par une action de développement, directement ou non, intentionnellement ou non. Se rapporte aux principaux impacts et effets de l'action sur les indicateurs locaux en matière de développement social, économique, environnemental et de développement en général. L'étude doit porter à la fois sur les résultats attendus et inattendus et doit également inclure les impacts positifs et négatifs de facteurs externes tels que l'évolution de l'opinion publique sur la corruption. Pour évaluer l'impact d'un programme ou d'un projet, il est utile de réfléchir aux questions suivantes :

- A quels résultats le programme ou le projet a-t-il abouti ?
- Quelle différence l'action a-t-elle réellement induit pour les bénéficiaires ?

La viabilité : Mesure si les bénéfices résultant d'une action sont susceptibles de se prolonger après la fin de l'intervention. Pour évaluer la viabilité d'un programme ou d'un projet, il est utile de réfléchir aux questions suivantes :

- Dans quelle mesure les bénéfices d'un programme ou d'un projet se prolongent-ils après le désengagement des donateurs ?
- Quels sont les facteurs qui ont favorisé ou empêché la viabilité d'un programme ou d'un projet ?

Source : Adapté de CAD/OCDE, 2011.

5.4 Les méthodes d'analyse pour les évaluations

Il existe tout un éventail de méthodes analytiques permettant de déterminer les réalisations et les impacts des interventions des AAC. Nous l'avons vu, les évaluations d'AAC entreprises jusqu'à présent font souvent l'erreur de présenter les évaluations de capacités organisationnelles comme la meilleure méthodologie de mesure des performances d'une agence, et n'emploient généralement que les entrevues avec des personnes ressources comme unique méthode d'analyse. Le fait que les AAC soient des organisations n'implique pas qu'il soit impossible de mesurer les résultats et les impacts de leurs interventions par le biais d'autres méthodes.

Le choix des méthodes d'analyse est déterminé par les questions de l'évaluation. Si l'évaluation porte sur le critère d'efficacité, l'évaluateur tentera d'identifier les activités de l'AAC et de mesurer les ressources internes utilisées et les produits réalisés, tâche qui n'implique pas de recourir à des techniques d'analyse très avancées. Cependant, si l'évaluation cible l'efficacité ou l'impact des interventions de l'AAC, l'évaluateur devra alors en mesurer les résultats et les impacts, ce qui exige des techniques d'analyse plus sophistiquées.

La présentation qui suit s'inspire du travail de McGee et Gaventa (2010), qui ont identifié plusieurs études basées sur des méthodes d'analyse permettant de démontrer soit la contribution soit l'attribution d'un impact.¹⁷ Les auteurs présentent quatre méthodes utilisées avec succès :

- *Les enquêtes* : Méthode quantitative pouvant servir de multiples objectifs, générant des données concrètes souvent appréciées des parties prenantes. Les Enquêtes de suivi des dépenses publiques (ESDP), par exemple, consistent à « pister l'argent » et à comparer les ressources reçues avec les dépenses effectuées afin de quantifier l'ampleur des fuites. D'autres enquêtes ne ciblent pas les éventuels écarts mais plutôt la perception des individus, comme c'est le cas des Enquêtes quantitatives sur la prestation de services (EQPS) et des Evaluations citoyennes (Reinikka et Svensson, 2000, 2003 ; Sundet, 2008 ; Ravindra, 2004). Les AAC peuvent gagner à utiliser ces méthodologies sophistiquées, mais n'ont pour autant pas à mettre en place des enquêtes complexes ou exhaustives pour parvenir à mesurer les résultats et les impacts de leurs interventions. Si les données collectées sont suffisamment complètes, les enquêtes quantitatives peuvent permettre une analyse statistique poussée (comme les analyses de régression à variables multiples ou les méthodologies de différence de différences utilisées dans les évaluations d'impact). D'autre part, des enquêtes plus qualitatives peuvent également s'avérer très utiles pour mesurer les performances d'une AAC.
- *Les démarches expérimentales* : Incluent les essais contrôlés randomisés, qui observent des principes scientifiques de base pour étayer l'analyse. De tels essais sont utilisés depuis des années en médecine et depuis peu en économie. Il n'est pas toujours possible, ni recommandé, d'employer des méthodes expérimentales en sciences sociales. Cependant, les données produites par le biais d'approches expérimentales ou quasi-expérimentales ont généralement plus de poids que les informations générées *via* d'autres méthodes. Une étude quasi-expérimentale est une méthode qui comporte en partie les caractéristiques des démarches expérimentales mais ne répond pas entièrement aux critères de validité internes (Bryman, 2004, 39). Olken (2007), ainsi que Bjorkman et Svensson (2009), ont par exemple mené des évaluations basées sur des méthodes expérimentales. Olken a ainsi pu examiner la différence des effets induits entre, d'un côté, des audits externes et, de l'autre, une surveillance de la communauté et un suivi sur le terrain en matière de prévention de la corruption autour de projets de routes dans 600 villages. Du fait de l'ampleur de l'échantillon évalué, l'étude a pu être randomisée, ce qui constitue un aspect essentiel des méthodes expérimentales. Ainsi,

¹⁷ Nick Duncan fournit également une présentation de méthodes d'analyse plus pointues (2006).

certaines projets furent aléatoirement sélectionnés pour faire l'objet d'un audit alors que d'autres furent soumis à des mécanismes de surveillance communautaire. Olken put ainsi intégrer les méthodes d'évaluation au programme dès sa phase initiale ; une approche expérimentale n'aurait autrement pas été réalisable. Olken a par ailleurs employé une méthode de stratification afin d'éviter tout effet de contamination ou de propagation.

- *Les études de cas qualitatives* : Souvent employées dans le domaine de la lutte contre la corruption. En effet, les évaluations qui ne s'appuient sur aucune méthodologie explicite sont souvent qualifiées d'études de cas. Les études de cas reposent cependant sur des règles méthodologiques qu'il convient de respecter (consulter, par exemple, Yin (2003) et Stake (1995)). Elles sont également faciles à associer avec d'autres méthodologies, ce qui permet d'en contextualiser et d'en approfondir les conclusions et d'expliquer d'éventuelles corrélations. Wade (1982, 1985) est l'un des premiers à avoir réalisé des études de cas très instructives.
- *Les interviews* : Méthode de collecte de données qualitatives populaire et très employée, souvent associée de manière systématique à d'autres méthodes (aussi bien qualitatives que quantitatives) afin d'obtenir une méthodologie de recherche combinée. Les interviews peuvent être structurées, semi-structurées ou libres. Pour les interviews collectives sur un sujet particulier, il est possible de recourir à des groupes témoins.

En général, il est préférable de procéder aux recherches par triangulation. Dans un domaine aussi complexe que la lutte contre la corruption, une évaluation qui repose sur une seule source de données ou technique d'analyse risque de générer des conclusions discutables. La triangulation permet de vérifier la crédibilité de ces dernières par recoupement. Les bonnes stratégies de recherche emploient souvent des méthodes à la fois qualitatives et quantitatives (méthodes mixtes), appliquées séquentiellement et selon des procédures itératives. Une stratégie de recherche itérative favorise une interaction continue entre la collecte et l'analyse des données ; autrement dit, l'analyse commence dès que les premières données ont été collectées, et les conclusions préliminaires qui en sont tirées déterminent l'étape de collecte de données suivante (Bryman, 2004, 399).

5.5 La participation

Trop souvent, les évaluations sont considérées comme des exercices externes et n'impliquent pas réellement les parties prenantes et les bénéficiaires. Les évaluations d'AAC que nous avons décrites dans le chapitre 2 n'échappent pas à la règle. L'aspect de la participation devrait être revalorisé, non pas comme une fin en soi, mais parce qu'il permet la validation des données et des conclusions des évaluations et est, à ce titre, vecteur de qualité. Par ailleurs, les bénéficiaires devant assumer seuls leurs responsabilités après le désengagement des donateurs, la participation garantit la viabilité des initiatives sur le long terme.

Les termes de référence d'une évaluation devraient favoriser autant que possible la participation du personnel des AAC, des donateurs et des parties prenantes externes à la procédure d'évaluation. Ces dernières peuvent être des ONG intervenant sur des questions de redevabilité ou encore, des organismes publics impliqués dans la lutte contre la corruption – autant d'intervenants qui sont en plus souvent susceptibles d'apporter des données intéressantes. L'opinion des bénéficiaires, c'est-à-dire de la population en général et, plus particulièrement, des personnes qui ont bénéficié des services de l'AAC, devrait également être sollicitée. Les évaluations d'AAC sont bien trop souvent laissées au « jugement des experts ». L'opinion et l'expertise des groupes cibles devraient pourtant faire partie intégrante des évaluations et ce, dès leur phase initiale. Ces derniers pourraient par exemple contribuer à l'élaboration des évaluations en identifiant les points à traiter et les questions d'évaluation à poser (OCDE, 2010a, 9, 23).

Les parties prenantes sont les individus ou les organisations susceptibles d'être affectés par les activités de l'AAC et/ou d'influencer le succès ou l'échec de ses interventions. Il existe différents instruments d'identification des parties prenantes, parmi lesquels le diagramme de Venn ou le graphique de type radar.¹⁸

La transparence est un autre facteur qui devrait être encouragé. Les commentaires des citoyens, des donateurs et des autres parties prenantes peuvent contribuer à la validité et à la fiabilité des conclusions des évaluations, c'est pourquoi ces dernières devraient être publiques.

5.6 La garantie de qualité

La question de la garantie de qualité devrait intéresser les AAC et le personnel des organismes donateurs chargé des évaluations d'AAC. Comme nous l'avons mentionné plus haut, en 2006, le Réseau du CAD/OCDE sur l'évaluation du développement a établi les « Normes du CAD pour une évaluation de qualité », dans lesquelles il fixe 10 critères de mesure de la qualité des évaluations (CAD/OCDE, 2006 ; consulter également CAD/OCDE, 2010b). Ces critères sont ceux que nous avons utilisés dans le cadre de l'inventaire d'évaluations que nous avons présenté dans le chapitre 2 : justification, finalité et objectifs de l'évaluation ; champ couvert par l'évaluation ; contexte ; méthodologie d'évaluation ; sources des informations ; indépendance ; déontologie de l'évaluation ; garantie de qualité ; pertinence des résultats de l'évaluation ; et exhaustivité.

À l'avenir, les AAC et le personnel des organismes donateurs devraient exiger des évaluateurs qu'ils respectent au moins ces critères, ou normes de garantie de qualité, et veiller à ce que leurs termes de référence reflètent cette exigence. La garantie de qualité commence en effet par des termes de référence bien conçus. Au cours de l'évaluation, il est important que l'AAC et/ou le responsable d'évaluation de l'organisme donateur effectuent une supervision constructive de la procédure. La première version du rapport doit faire l'objet d'une révision au titre de la procédure de garantie de qualité. Le plan d'évaluation est censé prévoir le délai nécessaire pour que le rapport soit transmis et que les commentaires et les réserves reçus en retour soient intégrés à la version finale, qui devrait en outre indiquer clairement s'il reste d'éventuelles divergences de vue. Les révisions par les pairs sont également facteurs de qualité. Les AAC devraient partager les conclusions de leurs évaluations, et les évaluateurs qui ont des doutes devraient consulter d'autres AAC.

5.7 Les aspects budgétaires et financiers du S&E

Comme nous l'avons indiqué dans le chapitre 2, les budgets alloués aux évaluations d'AAC sont souvent insuffisants pour couvrir une étude approfondie. Les donateurs doivent se montrer plus réalistes sur le coût des évaluations, et devraient surtout veiller à mieux optimiser leurs ressources.

Actuellement, les AAC font l'objet d'évaluations fréquentes mais superficielles. Les fonds seraient alloués à meilleur escient s'ils étaient consacrés à l'élaboration d'un système de S&E qui, conçu en collaboration avec le personnel des agences, serait capable d'effectuer des collectes régulières de données de suivi sur les indicateurs de produits, de réalisations et d'impact. Les coûts de conception et d'entretien d'un système de S&E devraient être inscrits au budget ordinaire ; par exemple, 5 à 10% du budget total pourraient être affectés au S&E. L'argent serait d'autant mieux dépensé que cela favoriserait une meilleure gestion axée sur les résultats et renforcerait les capacités des agences.

Comme nous l'avons mentionné plus haut, l'essentiel est que les systèmes de S&E restent simples et que des indicateurs et des enquêtes rapides soient identifiés pour permettre de mesurer résultats et

¹⁸ Pour plus d'informations sur l'analyse des parties prenantes, consulter la méthode en cinq étapes décrite dans les Lignes directrices Gestion du cycle de projet de EuropeAid (Commission européenne, 2004).

impacts. Une fois les systèmes de S&E en place, les donateurs devraient être en mesure d'entreprendre des évaluations externes plus conséquentes tous les cinq à sept ans. Cela libèrerait des fonds pour ce que l'on appelle les « évaluations d'impact », que nous allons maintenant traiter dans le chapitre 6.

6. La méthodologie d'évaluation d'impact

Sous la pression des donateurs, et face aux revendications croissantes des gouvernements nationaux et de la population, les AAC doivent rendre des comptes sur les impacts de leurs interventions. Le problème est que les agences ne disposent actuellement pas des instruments nécessaires pour satisfaire à ces demandes et démontrer leur valeur à la société par des preuves tangibles.¹⁹ Les évaluations d'impact, composantes centrales pour l'élaboration de politiques factuelles, peuvent constituer une source d'informations utiles aux stratégies de communication des AAC.

Les évaluations d'impact sont de plus en plus populaires, en particulier auprès des donateurs, car elles permettent d'effectuer des mesures au-delà du stade des produits. Jusqu'à présent, alors qu'ils ont investi massivement dans la mise en œuvre de programmes de développement, les donateurs et leurs partenaires locaux n'ont cependant rassemblé que peu de preuves concrètes de l'impact de leur soutien financier. Plusieurs raisons sont en cause, parmi lesquelles le peu d'intérêt et de financements accordés aux évaluations d'impact, mais également la qualité mitigée de celles qui ont été effectuées jusqu'à présent. Dans son rapport intitulé *When Will We Ever Learn?*, le Center for Global Development (2006) rend très bien compte de ces lacunes dans les secteurs sociaux.

La méthodologie d'évaluation d'impact n'a pour l'instant pas encore été appliquée dans le domaine de la lutte contre la corruption,²⁰ ce qui ne signifie pas pour autant que ce soit impossible. Les interventions anti-corruption ne sont en effet pas plus difficiles à mesurer que les actions menées dans les nombreux autres secteurs dans lesquels les évaluations d'impact sont fréquentes. Nous ne présenterons pas ici l'intégralité de la conception d'une évaluation d'impact, car la démarche est fonction de l'envergure souhaitée de l'évaluation et qu'il en existe par ailleurs autant de modèles que d'AAC. Nous traiterons toutefois les principaux éléments composant une évaluation d'impact d'AAC.

Nombre d'évaluations visent à mesurer des impacts, mais peu recourent en fait à une méthodologie d'évaluation rigoureuse qui en rende réellement compte. Une évaluation n'est pas tenue de procéder à une analyse d'impact exhaustive, mais les méthodes de recherche employées peuvent néanmoins gagner à s'inspirer des principes régissant les évaluations d'impact.

Bamberger (2006, 2–3) établit trois critères minimaux pour réaliser des évaluations d'impact de qualité :

- Elaborer un *ensemble d'indicateurs* permettant de définir et de mesurer de manière claire et fiable les moyens disponibles pour un projet, la mise en œuvre de ce dernier ainsi que ses produits, ses réalisations et ses impacts.
- Développer une analyse *contrefactuelle* solide, étayée par un raisonnement démontrant de façon convaincante que les variations enregistrées après l'intervention par les indicateurs de résultats sont bien dues au projet et non à d'autres éventuels facteurs sans rapport avec l'intervention tels qu'une amélioration de l'économie locale ou les programmes d'autres agences.

¹⁹ Une publication récente de l'Office des Nations unies contre la drogue et le crime et de la Banque mondiale étudie l'importance du soutien public pour les AAC et présente les outils de communication que ces dernières ont à leur disposition pour bénéficier de ce genre de soutien. Consulter Byrne, Arnold, et Nagano (2010).

²⁰ Il existe une multitude d'enquêtes générales appliquant des indicateurs « globaux » permettant de mesurer les variations enregistrées au niveau des impacts. Bien qu'elles fournissent des informations utiles, elles ne visent pas à démontrer les résultats d'interventions particulières. Ce ne sont donc pas des évaluations en tant que telles.

- Déterminer, *via* les procédures statistiques d'usage, *si un projet a contribué aux impacts escomptés* et profité à une bonne part de la population ciblée.

Nous adopterons ici une approche méthodologique pragmatique des évaluations d'impact. Nous partageons le point de vue des auteurs selon lesquels « il convient d'identifier les méthodes les plus appropriées au contexte opérationnel, et non pas le contraire » (Gertler et al., 2011, xiii–xiv). A cet effet, l'approche la mieux indiquée est l'approche prospective, qui intègre la méthodologie de l'évaluation à celle de la mise en œuvre du programme.

6.1 Les composantes du modèle d'évaluation d'impact d'une AAC

Les évaluations d'AAC ne peuvent pas toujours s'appuyer sur la meilleure des méthodologies statistiques de recherche et se baser sur la randomisation ou l'appariement statistique. Cela dépend à chaque fois si les évaluateurs ont la possibilité d'entreprendre une évaluation prospective et de compter sur des données de qualité. Dans la plupart des cas, le projet doit s'appuyer sur une méthode quasi-expérimentale ou non expérimentale. D'autres aspects de la méthodologie de l'évaluation peuvent toutefois venir compenser cette faiblesse.

Les spécialistes divergent sur la question et se demandent si le manque de données statistiques peut être entièrement compensé par d'autres points forts de la méthodologie d'évaluation au point de générer une réelle « évaluation d'impact ». Certains spécialistes, appelés les « randomistas », argumentent que les évaluations d'impact devraient systématiquement se baser sur une méthodologie expérimentale aléatoire. D'autres considèrent que des méthodes quasi-expérimentales bien conçues peuvent constituer une seconde option tout à fait acceptable. Certains auteurs estiment également que pour être « satisfaisante », une évaluation d'impact doit reposer sur des méthodes d'évaluation contrefactuelles étayées par une analyse de groupes de comparaison statistiquement appariés. Enfin, d'autres sont ouverts à l'utilisation de bases alternatives aux méthodes contrefactuelles, comme les données secondaires, les modèles de théories de programmes (théories du changement), les analyses de résultats ou les schématisations conceptuelles. La plupart des praticiens de la lutte contre la corruption s'accordent cependant à penser que ce débat reste purement théorique et que, dans la pratique, il est plus important de définir les moyens de mettre en œuvre la meilleure méthode praticable possible pour mesurer l'impact des projets anti-corruption.

ENCADRE 8. LES CRITERES D'EVALUATION EN MATIERE DE RECHERCHE SOCIALE

La fiabilité, la reproductibilité et la validité sont les trois critères les plus importants pour l'évaluation en matière de recherche sociale :

- *La fiabilité* renvoie à la reproductibilité des résultats d'une étude et à la cohérence des mesures effectuées.
- *La reproductibilité* est étroitement liée à la fiabilité. Si le chercheur explique clairement sa méthode, l'étude aura plus de chances d'être reproductible.
- *La validité* porte sur l'intégrité des conclusions de la recherche. Ce point peut être divisé en (a) la validité des mesures, (b) la validité interne, (c) la validité externe, et (d) la validité environnementale.

Source: Bryman, 2004, 28.

Indépendamment de la possibilité de recourir à la randomisation ou à l'appariement statistique, nous préconisons trois méthodes sont susceptibles de permettre l'établissement d'un protocole sérieux d'évaluation d'AAC :

- Baser l'évaluation sur une *théorie du changement* solide, à partir de laquelle établir des indicateurs, des hypothèses de recherche et des questions ;

- Employer un protocole d'évaluation à *méthodes mixtes* ; et
- Utiliser *un grand nombre d'indicateurs variés, aussi bien au niveau des résultats que des impacts*, afin d'améliorer la validité conceptuelle de l'évaluation, en décomposant les concepts complexes tels que la corruption en analyses plus détaillées (Bamberger, Rugh, et Mabry, 2006 ; Bamberger et White, 2007).

Voyons maintenant ces trois points plus en détail.

6.1.1 La théorie du changement

Au cours des chapitres précédents, nous avons vu les principales caractéristiques de la logique de la théorie du changement. Une évaluation qui repose sur une théorie, c'est-à-dire un « examen des hypothèses de base des liens de causalité entre, d'un côté, les moyens et, de l'autre, les résultats et les impacts », constitue une méthode bien établie défendue, entre autres, par White (2009). Mesurer l'impact d'une intervention par le biais d'une méthode théorique revient à analyser non seulement ce qui fonctionne et pourquoi, même lorsque le protocole employé ne prévoit pas de randomisation ni d'expérimentation. Les évaluations d'impact basées sur une théorie sont donc un instrument adéquat pour mesurer l'impact des interventions d'AAC, pour lesquelles il est souvent impossible de disposer de groupes de comparaison ou de contrôle.²¹

Alors qu'il effectuait une comparaison de trois différents types d'interventions ordinaires, Ravallion (2009) recommanda d'utiliser une méthode basée sur une théorie pour évaluer l'impact des commissions anti-corruption. Il proposa notamment de planifier une couverture géographique ou ministérielle des interventions de l'AAC (par exemple, un système de dénonciation), de manière à pouvoir en mesurer l'impact sur un ministère ou un service de police local particulier par rapport à ceux qui ne bénéficient d'aucune intervention. Ces interventions peuvent ensuite faire office de scénarii contrefactuels.²²

Comme le savent la plupart des praticiens anti-corruption, les explications théoriques de la corruption peuvent varier fortement. Il est donc fondamental de parvenir à instaurer un consensus entre la multitude de parties prenantes sur les théories du changement d'un programme. Cela ne devrait pas poser trop de problèmes, dans la mesure où les théories du changement ne sont pas des vérités établies sinon une cristallisation de la logique implicite d'un programme. C'est cette logique de causalité qui devrait être testée.

Pour citer un exemple simple de liens de causalité, le financement d'un donateur (moyens) permet à une AAC de mettre en service un système de traitement des plaintes en ligne (activités) qui donne lieu à x enquêtes (produits) pouvant aboutir à x condamnations (réalisations) et, à terme, contribuer à réduire le taux de corruption (impact). Il s'agit bien sûr d'un modèle simplifié ; l'intégralité des liens de causalité recouvre de nombreux autres éléments.

²¹ D'après Martin Ravallion (2009, 227–36), les AAC ne sont pas exactement des « programmes assignés » : elles ne ciblent généralement pas uniquement certaines unités d'observations, elles ont plutôt une portée nationale. Il est par conséquent plus difficile d'établir des groupes témoins dans ce contexte.

²² Un biais de sélection risque de donner lieu à une évaluation partielle, car les unités de traitement seraient probablement sélectionnées en raison de la prévalence de la corruption et différeraient donc du groupe de contrôle. Différentes solutions existent toutefois qui permettent de résoudre ce problème.

6.1.2 Les méthodes mixtes

Un protocole de méthodes mixtes repose sur l'utilisation itérative et complémentaire de méthodes d'analyse à la fois quantitatives et qualitatives. Chaque méthode ayant ses forces et ses faiblesses, les associer permet d'optimiser les résultats des évaluations.

Les protocoles de méthodes mixtes permettent de résoudre un grand nombre des problèmes qui entachent habituellement la validité des évaluations statistiques d'impact. Parmi leurs principaux avantages, on peut citer :

- *Des mesures plus précises, et donc une validité conceptuelle renforcée* : L'utilisation de protocoles de méthodes mixtes, tout comme le recours à de multiples indicateurs, contribue à renforcer la validité conceptuelle d'une évaluation. La validité conceptuelle renvoie à la capacité d'une évaluation à mesurer la variable qu'elle est censée mesurer. (Par exemple, dans quelle mesure un test de QI est-il capable de mesurer l'intelligence est une question classique de validité conceptuelle). La corruption est un concept social complexe ; il est important que les évaluations en mesurent bien les différents aspects. L'étude de la corruption repose souvent sur des variables supplétives qui ne couvrent pas entièrement le concept analysé (la corruption). En recourant à des méthodes qualitatives telles que les études de cas exploratoires, les groupes témoins ou les interviews, l'évaluateur peut aider à mieux comprendre la variable clé de l'évaluation, à savoir, la corruption.
- *La contextualisation de l'analyse* : Les protocoles purement quantitatifs d'évaluation statistique d'impact peuvent jouir d'une bonne validité interne, mais ils risquent de passer à côté de facteurs politiques, institutionnels, historiques et socioculturels susceptibles d'avoir influencé positivement ou négativement les interventions. Intégrer une analyse économique-politique et contextuelle au protocole d'évaluation permet de clarifier la théorie du changement utilisée et contribue donc à crédibiliser les conclusions de l'évaluation. Identifier l'impact d'une intervention exige de bien comprendre le cadre social, politique et économique dans lequel cette dernière s'inscrit. Par exemple, la mise en place d'un système de traitement des plaintes demandera de tenir compte non seulement de questions de développement (par exemple, du niveau des connaissances informatiques et de l'accès à Internet) mais également d'enjeux techniques (par exemple, de la nécessité de sécuriser l'identité des utilisateurs) et de questions culturelles (par exemple, de la tendance à ne pas signaler de comportements malhonnêtes par peur de représailles). En outre, certaines questions d'ordre juridique (par exemple, l'existence et la mise en vigueur d'une législation de protection des dénonciateurs) exerceront un effet externe sur le niveau des plaintes enregistrées.
- *Le problème de la « boîte noire » et la procédure de mise en œuvre du projet* : La plupart des évaluations quantitatives d'impact visent uniquement à déterminer si un projet a atteint ses objectifs, et comparent pour cela les niveaux avant et après l'intervention. Cette méthode ne permet pas de définir si un échec est dû à l'application d'un mauvais protocole ou à une mauvaise mise en œuvre du projet. Des techniques qualitatives peuvent rendre une image fidèle des procédures appliquées au projet et prévoir de recouper ces informations avec des données quantitatives.
- *L'usage fréquent de la triangulation* : Originellement conceptualisée par Webb et al. (1966), la triangulation implique de recourir à plusieurs méthodes ou sources de données pour étudier des phénomènes sociaux. Denzin (1970, 310) emploie des termes plus généraux et désigne une méthode associant « un grand nombre d'observateurs, d'approches théoriques, de sources de données et de méthodologies. » L'emploi simultané de méthodes quantitatives et qualitatives permet à l'évaluateur de mesurer la validité des

différentes sources de données en effectuant des comparaisons systématiques de ces dernières les unes avec les autres et en procédant à des contrôles de cohérence (Bamberger, Rao et Woolcock, 2010).

En somme, les évaluations d'AAC ne peuvent que bénéficier de l'utilisation de méthodes mixtes. Pour l'instant, les approches les plus employées dans ce domaine sont qualitatives (interviews, analyses documentaires, etc.), en partie parce que les AAC n'ont pas élaboré de systèmes de S&E axés sur les résultats pour collecter leurs données de référence et leurs données systématiques. Les agences doivent consacrer plus de temps et de ressources à l'établissement d'approches quantitatives, qu'elles doivent associer à des méthodes qualitatives.

6.1.3 Des indicateurs variés à tous les niveaux

Les praticiens de la lutte contre la corruption ont particulièrement tendance à s'appuyer sur un seul et unique indicateur ou indice pour rendre compte des impacts des interventions. Généralement, ils ont recours à des indices de niveau macro tels que l'IPC de Transparency International ou les IMG de la Banque mondiale, popularisés dans les années 1990 par la « révolution des données ». Il est toutefois erroné de considérer un phénomène aussi complexe que la corruption à travers un seul indicateur ou indice. Il est impossible de déduire qu'une institution a à elle seule un quelconque effet par la seule mesure d'un résultat de niveau macro. Comme nous l'avons mentionné plus haut, les indicateurs devraient être décomposés au maximum afin d'affiner autant que possible la mesure du concept.

La demande pour les mesures d'impact générales reste cependant constante. Nombre des indices de corruption ne sont pourtant pas toujours cohérents entre eux, en partie en raison du manque de validité des méthodologies employées. Différents spécialistes ont signalé les pièges que recèlent de tels indices, mais la popularité de ces derniers n'a pour autant pas été entamée (Galtung, 2005 ; Søreide, 2003, 2006 ; Arndt et Oman, 2006 ; Knack, 2006).

Lorsque les ressources le permettent, il est par ailleurs conseillé de recouper les données en complétant les indices génériques de niveau macro, qui saisissent rarement l'intégralité de la logique du programme, avec les indicateurs spécialement conçus aux fins de l'évaluation. Idéalement, les indicateurs de perception devraient être associés à des indicateurs « objectifs » basés sur des faits ou des expériences réelles (voir chapitre 7). Les indicateurs devraient refléter le fait que des résultats et des impacts positifs peuvent traduire non seulement une réduction de la corruption et de la criminalité, mais également une amélioration des relations avec les organisations partenaires, une hausse de la sensibilisation du public, une valorisation des compétences, etc. Mais on ne peut attendre d'une AAC qui n'a aucun pouvoir d'enquête ou de poursuites qu'elle contribue à faire évoluer les statistiques criminelles, par exemple. Même les AAC qui disposent de pouvoirs d'enquête sont peu concernées par la mesure quantitative des impacts, car les grandes affaires complexes de corruption peuvent atteindre une telle envergure que des mesures numériques telles que le nombre d'affaires ayant donné lieu à des poursuites n'ont ici aucun intérêt. Dans le chapitre 7, nous traiterons la question des indicateurs et verrons leur importance vis-à-vis du suivi et de l'évaluation.

6.1.4 Résumé

Les évaluateurs devraient toujours s'évertuer à employer le meilleur protocole d'évaluation possible. Ils sont toutefois confrontés à un certain nombre d'obstacles et de limites qui les contraignent à improviser, s'adapter et aller de l'avant. Même lorsqu'il est impossible d'appliquer un protocole expérimental statistique, les évaluations d'impact des AAC peuvent se baser sur des protocoles quasi-expérimentaux ou non expérimentaux du moment qu'ils soient étayés par au moins trois piliers méthodologiques : (a) une théorie du changement solide, (b) des méthodes mixtes, et (c) des indicateurs variés à tous les niveaux.

En fait, même si certains obstacles méthodologiques limitent le potentiel des évaluations, les donateurs n'ont aucune raison de ne pas promouvoir les évaluations d'impact dans le domaine de la lutte contre la corruption, et plus particulièrement des AAC. Les lacunes statistiques peuvent être compensées en renforçant le protocole d'évaluation sur d'autres fronts.

Nous consacrerons la suite de ce chapitre à des questions méthodologiques spécifiques aux évaluations d'impact, qui intéresseront tout particulièrement les évaluateurs expérimentés. Nous aborderons tout d'abord les pratiques actuellement en vigueur en matière d'évaluation d'impact, puis traiterons la question de l'attribution. Enfin, nous présenterons différentes méthodes d'élaboration de scénarii contrefactuels pour les interventions des AAC.

6.2 Les pratiques actuelles en matière d'évaluation de l'impact des AAC

Actuellement, les donateurs et les évaluateurs recourent essentiellement à deux méthodes pour mesurer l'impact des interventions anti-corruption. Les deux sont trompeuses. Tout d'abord, parce que certaines évaluations sont appelées « évaluations d'impact » alors qu'elles ne mesurent en réalité aucun impact particulier. Ensuite, parce que les « indicateurs globaux », les indices transnationaux, les classements internationaux et les autres instruments du genre sont parfois utilisés pour illustrer des impacts alors qu'aucune corrélation réelle ne puisse être démontrée.

En ce qui concerne le premier point, il arrive qu'une évaluation entreprise dans le but de mesurer un impact repose sur un protocole qui ne prévoit pas d'effectuer de telles mesures. Les donateurs ne doivent pas céder à la tentation de faire réaliser de telles « évaluations d'impact » s'ils n'envisagent pas de fournir aux évaluateurs les ressources plutôt substantielles qui leur seront nécessaires pour cette tâche (et non pas uniquement les ressources nécessaires aux évaluations organisationnelles, plus modestes), de même qu'ils ne devraient pas se fier à ceux qui affirment pouvoir exécuter un tel travail sans les ressources adéquates. Aucune des évaluations passées en revue dans le chapitre 2, par exemple, ne peut être considérée comme une évaluation d'impact.

L'examen des évaluations que nous avons effectué fait par ailleurs ressortir que les donateurs tendent à fixer un délai d'évaluation beaucoup trop court. Il y aura toujours un décalage entre le début de l'intervention et le moment où l'on peut raisonnablement espérer observer des résultats et, à plus forte raison, des impacts. Par conséquent, entreprendre une évaluation d'impact d'une AAC deux ans après sa création ne suffira pas à déterminer le succès de cette dernière, mais donnera seulement une première indication de ses performances. Cela ne signifie pas pour autant qu'il soit inutile d'évaluer l'impact des interventions d'AAC, mais seulement que de telles évaluations doivent à la fois être entreprises de manière rigoureuse, se baser sur des faits concrets et sur une bonne connaissance du mandat, des fonctions et du contexte de l'agence et enfin, s'inscrire dans un délai raisonnable.

Le deuxième point, qui concerne les indicateurs globaux, est plus complexe. Il est également lié à la question de la redevabilité et de l'attribution que nous aborderons dans le chapitre 6.3. Les indicateurs globaux, les enquêtes transnationales et les indices internationaux sont facilement accessibles. Ils sont en plus instructifs et ciblent bien le niveau des impacts (comme par exemple le niveau perçu de corruption dans un pays). Les employer pour évaluer dans quelle mesure une AAC a atteint ses objectifs n'est pour autant pas indiqué, même si l'AAC en question vise par exemple la réduction de la corruption ou l'évolution des valeurs dans une société. Le problème est qu'il est « pratiquement impossible d'identifier des corrélations entre les AAC et des résultats généraux de ce genre, notamment parce qu'il est nécessaire de rendre compte de l'influence des réformes structurelles et d'autres facteurs importants. Il faut donc se méfier des affirmations simplistes concernant l'impact d'une AAC sur la corruption. La mission des agences est en effet bien souvent définie en des termes très généraux de réduction de la corruption ou de modification des valeurs – des résultats qui sont, au mieux, extrêmement difficiles à mesurer » (USAID, 2006, 4). L'utilisation exclusive d'indicateurs globaux génériques (associés avec différentes appréciations des produits) n'est effectivement pas suffisante pour identifier les corrélations en jeu. En revanche, la méthodologie proposée dans le

présent chapitre (c'est-à-dire l'application d'une théorie du changement, de méthodes mixtes et d'indicateurs variés à tous les niveaux) établit ces liens, de même qu'elle permet de mesurer les impacts et les réalisations des interventions effectuées.

Pour résumer, il est impossible de conclure, sur la seule base d'indicateurs généraux, qu'un projet ou une institution a un impact particulier. Pour mesurer l'impact du travail d'une AAC, l'évaluation doit être considérée dans la perspective du programme qui est mis en œuvre. En somme, les indicateurs globaux ne sauraient constituer des raccourcis pour mesurer l'impact d'une AAC.

6.3 La redevabilité et l'attribution

Réduire ou, tout au moins, réajuster l'écart qui existe entre les attentes et la réalité concernant certaines AAC nécessite une solide compréhension de la méthodologie d'évaluation d'impact. Alors qu'on attend souvent beaucoup des AAC, ces dernières ne disposent pas toujours des moyens, de la marge de manœuvre ou du temps nécessaires pour répondre à ces attentes. Au niveau des résultats et des impacts, les effets directs qui sont supposés être induits par les interventions des AAC sont probablement irréalistes. Dans la plupart des sphères de la lutte contre la corruption, les chercheurs n'ont même pas encore résolu les problèmes de liens de causalité généraux. Il est acquis qu'il existe une corrélation entre la transparence et la baisse de la corruption ; on ignore cependant si la transparence entraîne une baisse de la corruption ou si les mesures en faveur de la transparence ont tout simplement plus de chance d'être appliquées lorsque le niveau de corruption est moindre. Si, alors que les AAC sont incitées à promouvoir des mesures en faveur de la transparence, on ne peut démontrer que cela contribue directement à réduire le niveau de corruption, la faute ne leur en revient pas nécessairement. Attribuer des variations observées à une intervention et affirmer qu'il existe des liens de causalité entre les différents facteurs en jeu requiert une bonne maîtrise du débat théorique en la matière. A leur tour, les évaluations d'impact qui reposent sur une théorie peuvent sensiblement enrichir le débat théorique.

Généralement, il convient de commencer par déterminer si l'on observe des variations puis, le cas échéant, si ces variations peuvent être attribuées à l'intervention évaluée. Sur le plan méthodologique, les questions relatives aux liens de causalité, aux impacts et à l'attribution sont plus délicates, et y répondre exige d'élaborer un scénario contrefactuel. Lorsque les variations observées ne peuvent être attribuées à l'intervention, il est alors possible d'entreprendre une analyse de contribution. Plutôt que d'essayer d'évaluer la part de variations qui résulte de l'intervention (attribution), une analyse de contribution vise uniquement à démontrer si l'intervention évaluée est l'une des causes des variations observées. Les contributions peuvent être classées mais ne sont pas quantifiables. L'analyse est progressive et se base sur une chaîne de raisonnements logiques intégrant d'autres explications possibles et tous les facteurs externes en jeu.

Concevoir un solide cadre de gestion axé sur les résultats, une théorie du changement ou un cadre logique peut aider à déterminer le niveau de performance que l'organisation est susceptible d'atteindre en fonction de ses ressources et de ses contraintes. Une AAC qui dispose de ressources suffisantes doit pouvoir réaliser les produits visés et rendre des comptes sur la question. Elle n'est en revanche pas entièrement responsable de ses réalisations. Si elle suit un modèle

ENCADRE 9. L'ANALYSE D'ATTRIBUTION ET L'ANALYSE DE CONTRIBUTION

L'analyse d'attribution vise à évaluer la part des variations observées pouvant être réellement attribuée à une intervention donnée. Cela implique d'établir un scénario contrefactuel, c'est-à-dire dans lequel l'intervention n'a pas eu lieu.

L'analyse de contribution vise à démontrer si l'intervention évaluée contribue bien aux variations observées. Peut également classer l'intervention évaluée parmi les différentes causes des variations observées.

Source: Commission européenne, 2006a.

logique valable, les produits devraient aboutir aux réalisations souhaitées, mais des facteurs externes sont susceptibles d'interférer. Le niveau des réalisations constitue donc une zone grise en matière de redevabilité, car les AAC n'en sont responsables qu'en partie seulement.

Au niveau le plus élevé, celui des impacts, qui est souvent le niveau de l'objectif global dans les cadres logiques, les AAC ne sont pas directement responsables de leurs impacts. On ne peut normalement pas espérer pouvoir attribuer la réalisation d'un objectif aux activités d'une AAC, mais uniquement que les réalisations de cette dernière contribuent de manière positive à l'objectif en question. Dans la pratique, cela a des conséquences importantes. Par exemple, la baisse de la corruption devrait-elle constituer un indicateur d'impact ou de résultat ? Tout dépend du type d'AAC concerné. Si la finalité de toutes les agences est de réduire la corruption, les AAC dont le rôle est préventif auront certainement moins d'influence sur le niveau de corruption, excepté par le biais indirect d'activités d'éducation et de sensibilisation du public. La diminution de la corruption ne constitue alors pas un résultat potentiel, mais uniquement un objectif global auquel l'AAC devrait normalement contribuer sur le long terme.

Deux concepts distincts sont importants dans ce contexte : celui de sphère d'influence et celui de sphère d'intérêt. La sphère d'intérêt des AAC est très vaste et s'étend jusqu'à la réduction de la corruption. Leur sphère d'influence, en revanche, n'est pas toujours aussi étendue.²³

Prenons un exemple. Les AAC sont souvent chargées de mettre en œuvre la stratégie nationale de lutte contre la corruption. Les indicateurs de la stratégie nationale et de l'AAC sont donc censés coïncider. La stratégie nationale ne dépend pas que de l'AAC mais également d'autres acteurs, c'est pourquoi il est important de distinguer son succès de celui de l'agence. Ainsi, tous les éléments d'une stratégie anti-corruption entrent normalement dans la *sphère d'intérêt (sphere of concern)* des AAC, mais certains échappent sans doute à sa *sphère de responsabilités (sphere of responsibility)*. Tous les acteurs impliqués dans la stratégie nationale de lutte contre la corruption *contribuent* à son objectif global (en général, la réduction de la corruption) mais, à ce niveau, aucun ne peut être directement tenu pour *responsable* de son échec ni se voir *attribuer* son succès. Il faudrait pour cela que l'objectif entre dans la sphère de responsabilités de l'AAC concernée et que cette dernière dispose des ressources nécessaires pour influencer directement les résultats visés. Cela demande généralement de cibler non pas l'objectif global, mais plusieurs sous-objectifs plus précis (consulter le catalogue d'indicateurs en annexe 1).

Le concept « d'association la plus probable » de la Banque mondiale est une variante de l'attribution, plus facile à manier pour les évaluateurs.²⁴ Les défenseurs de ces méthodologies argumentent qu'elles saisissent mieux la complexité des réalités que d'autres méthodes plus traditionnelles, qui tentent d'adapter « de force » la réalité à leurs paramètres.²⁵

²³ Ces concepts s'utilisent dans le cadre d'une méthodologie d'évaluation spécialisée appelée cartographie des incidences (outcome mapping), qui se base sur le fait que « la complexité du processus de développement rend l'évaluation des impacts extrêmement difficile (en particulier pour un organisme donateur externe qui tente d'effectuer une analyse d'attribution). » Cette méthodologie vise à mesurer les résultats et les changements de comportement des bénéficiaires directs plutôt que les impacts (Earl, Carden et Smutylo, 2001, 1–14). Note : Ressource en français : Earl, S., Carden, F., Smutylo, T. 2002. [Intégrer l'apprentissage et la réflexion dans les programmes de développement](#), IDRC, Canada.

²⁴ D'après les auteurs du concept « most likely association », il « permet une analyse éclairée qui, tout en se basant sur les meilleures données factuelles disponibles, admet que les conditions sont loin d'être expérimentales et que les lacunes sont nombreuses, tant en matière de données que de connaissances » (Foresti et al., 2007, 23).

²⁵ Les lecteurs intéressés par les méthodologies qui rendent compte des changements sociaux souhaiteront peut-être se documenter sur la cartographie des incidences (voir note 21) et sur la technique du « changement le plus significatif », méthode plus narrative que les méthodes quantitatives traditionnelles utilisées pour les analyses

6.4 Le principe de l'analyse contrefactuelle

L'évaluation d'impact est une méthode contrefactuelle, qui permet de procéder par élimination pour déterminer si c'est bien l'intervention de l'AAC évaluée, et non pas d'autres facteurs éventuels, qui est à l'origine des variations enregistrées. Pour tester et éliminer les autres explications possibles, il est essentiel d'établir des scénarii contrefactuels en parallèle avec les mesures de référence.

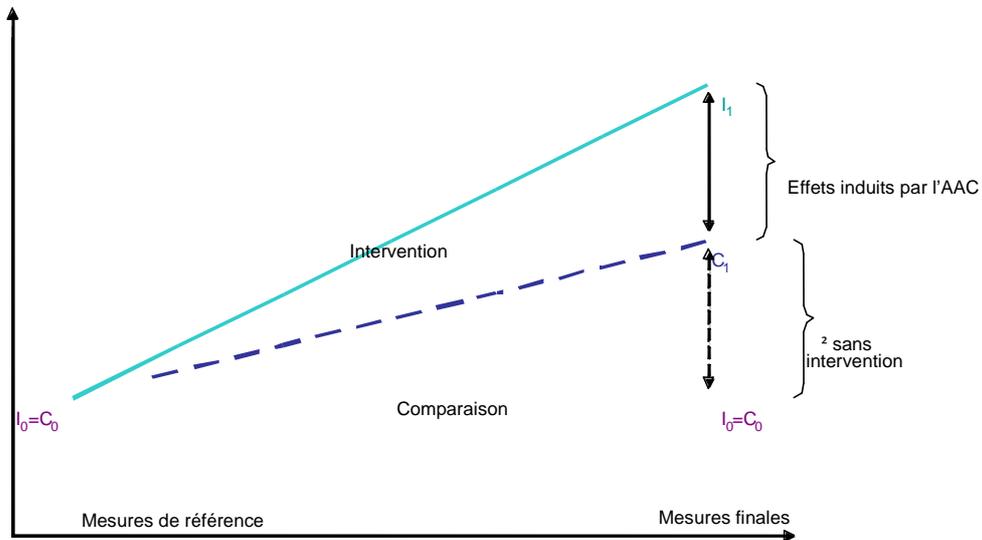
Elaborer des scénarii contrefactuels et des mesures de référence demande beaucoup de temps. Plutôt que de mesurer toutes les activités et les produits d'une AAC, il est donc judicieux de se limiter dès le début à n'identifier qu'un petit nombre d'indicateurs de résultats à suivre et, si possible, d'instituer des mesures pour un scénario contrefactuel les incluant.

Le recours à l'analyse contrefactuelle, par l'étude d'un groupe ou d'une entité de comparaison qui n'a pas bénéficié de l'intervention de l'AAC, est un élément central dans n'importe quelle évaluation d'impact. Comme le montre le schéma 4, pour réunir des données tangibles à propos de l'impact d'une AAC, il est nécessaire de trouver un groupe de comparaison illustrant un scénario contrefactuel, c'est-à-dire un groupe qui n'a pas été visé par l'intervention. On peut ensuite estimer l'impact net (score de différence) en comparant le groupe de traitement avec le groupe de contrôle, et isoler l'effet induit par l'intervention de l'AAC en retranchant la valeur du groupe de comparaison. Une simple analyse avant-après permet de conclure que l'AAC a bien produit les effets II. Ces conclusions sont toutefois trompeuses, dans la mesure où d'autres facteurs externes sont également susceptibles d'avoir des effets, en raison desquels le groupe de comparaison atteint C1. L'impact (valeur ajoutée) de l'AAC se mesure donc en retranchant C1 de II.

Dans la mesure où les AAC visent généralement des groupes cibles nationaux, elles peinent souvent à identifier des groupes cibles spécifiques et donc à définir des groupes n'ayant pas bénéficié de l'intervention pour le scénario contrefactuel. Ici encore, la décomposition des objectifs et des activités de l'agence (par exemple, en activités sectorielles) favorise la création de scénarii contrefactuels. Lorsque le scénario contrefactuel ne peut être illustré par un groupe de comparaison, il est alors possible de recourir au statu quo (absence d'intervention).

Le schéma 4 décrit une méthode d'évaluation d'impact stylisée. Il montre comment l'effet induit (« net out ») par l'AAC est déterminé en établissant la valeur ajoutée produite comparée à la situation si l'intervention n'avait pas eu lieu.

d'attribution (Davies et Dart, 2005). Cette méthodologie, ainsi que de nombreux autres documents relatifs au S&E, peuvent être consultés en anglais sur le site Internet de Monitoring and Evaluation News (<http://www.mande.co.uk>).

Schéma 4. Méthode d'évaluation d'impact stylisée

I_0 : Mesures de référence du groupe d'intervention
 C_0 : Mesures de référence du groupe de comparaison

I_1 : Mesures finales du groupe d'intervention
 C_1 : Mesures finales du groupe de comparaison

Toute évaluation d'impact rigoureuse dispose d'un groupe de comparaison. Dans le cas contraire, il est toutefois possible d'élaborer un scénario contrefactuel en imaginant les effets probables qui auraient été induits par l'absence d'intervention.

7. Les indicateurs de suivi des performances : Introduction au catalogue d'indicateurs

Elaborer des indicateurs requiert de bien cerner le phénomène de la corruption et de maîtriser les méthodes individuelles et institutionnelles destinées à l'enrayer. Nous avons ici consacré beaucoup d'efforts à l'élaboration d'indicateurs, car si ces derniers ne sont pas correctement adaptés à la réalité, même une méthodologie de recherche sérieuse et une exécution réfléchie de l'évaluation resteront vaines. La conception de bons indicateurs fait certes partie des responsabilités des AAC au titre de leur plan stratégique, mais évaluateurs et donateurs peuvent jouer un rôle constructif en leur proposant conseils, garantie de qualité et critiques.

Il est vrai que, dans le domaine de la lutte contre la corruption, les instruments de mesure ne sont pas aussi développés que dans d'autres secteurs comme la santé et l'éducation. Il existe cependant toute une batterie d'indicateurs utiles qui, associés entre eux, fournissent une image claire des performances d'une AAC. Utiliser des ensembles d'instruments ou d'indicateurs uniformes n'est pas une bonne solution, car le phénomène de la corruption est par nature fortement dépendant du contexte dans lequel il s'inscrit. Pour cette raison, évaluer les progrès réalisés dans la lutte contre la corruption constitue une véritable gageure. Le catalogue d'indicateurs que nous présentons en annexe 1 propose toute une gamme d'indicateurs possibles, qui sont à la fois souples et peuvent être adaptés en fonction des besoins.

7.1 L'élaboration d'indicateurs

Sans indicateurs, il est impossible de mesurer le succès d'une intervention. Un indicateur peut se définir comme une mesure systématiquement suivie dans le temps, qui traduit des variations positives ou négatives, ou l'absence de variations, dans les progrès réalisés pour atteindre une cible donnée. Pour évaluer si une intervention a des effets positifs, les dirigeants et les décideurs politiques devraient baser leur analyse sur l'ensemble des données tangibles générées par un groupe d'indicateurs. Aucun résultat, aucun impact ne devrait être mesuré à la lumière d'un seul indicateur. Généralement, les indicateurs proviennent de l'impact, des réalisations et des produits définis à l'avance comme performances visées. Chacun d'eux doit donc être soigneusement établi, tout comme la logique ou la théorie du changement de chaque intervention.

Les indicateurs peuvent servir aussi bien à des fins de suivi que d'évaluation, mais doivent faire l'objet d'un traitement différent suivant leur emploi. Une évaluation peut recourir aux mêmes indicateurs que ceux qui servent au suivi, mais elle doit adopter une approche plus large, donner lieu à une analyse plus approfondie, reposer sur d'autres sources de données et suivre une méthodologie adaptée à l'intervention. Les données de suivi sont néanmoins souvent

capitales pour effectuer une bonne évaluation, en particulier lorsqu'il s'agit d'une évaluation d'impact (Morra Imas et Rist, 2009, 108). Dans le catalogue proposé en annexe, nous présentons des indicateurs applicables à des fins de suivi, mais qui peuvent également être utiles aux évaluateurs.

ENCADRE 10. INDICATEURS, CONCEPTS ET MESURES

Un *concept* est une notion ou une idée générale utilisée pour définir un résultat. La bonne gouvernance, la qualité environnementale, etc., sont par exemple des concepts. Un *indicateur* représente un aspect clé du concept ; il est « indicatif » du concept, bien qu'il n'en constitue pas une représentation intégrale. Une *mesure* exprime la valeur quantitative ou qualitative d'un indicateur. Le mot indicateur est souvent employé à la place du mot mesure ; nous les distinguons pour mieux signifier que les *indicateurs* doivent être *mesurables*.

Les indicateurs doivent être le plus détaillés possible, de manière à pouvoir saisir les nuances, par exemple, des différents types de corruption, de la corruption sectorielle (dans les secteurs public et privé, la police, les douanes, etc.), entre les sexes, selon les lieux ou les méthodes de signalement de la corruption (par e-mail ou par courrier, en personne, etc.). Pour illustrer ce point, nous verrons plus loin l'exemple de la Commission indépendante contre la corruption (ICAC) de Hong Kong.

Nous préconisons l'usage de méthodes mixtes, aussi bien dans le sens traditionnel de l'association d'indicateurs quantitatifs et qualitatifs (par exemple, des statistiques quantitatives et des études de cas qualitatives) que dans le sens plus général d'association d'indicateurs basés sur la perception ou d'approximations avec des indicateurs plus « tangibles » qui mesurent directement la corruption (voir ci-dessous). Les méthodes mixtes ont plusieurs avantages. Elles permettent notamment de mieux appréhender le phénomène étudié, de recouper les sources d'information et de limiter les écarts susceptibles de découler d'une source de données unique.

Les indicateurs peuvent être classés en fonction de leur objectif et du type de mesures sur lesquelles ils reposent.

Dans la publication *Un guide de l'utilisateur pour mesurer la corruption* (2008, 9), le PNUD divise les indicateurs de corruption en cinq grands types :

- *Les indicateurs objectifs* : Indicateurs construits à partir de faits incontestés. Des exemples typiques pourraient inclure l'existence de lois anti-corruption ou le financement reçu par l'agence anti-corruption ou encore, des statistiques nationales.
- *Les indicateurs fondés sur la perception* : Indicateurs basés sur les opinions et les perceptions de corruption dans un pays donné émises par les citoyens et les experts.
- *Les indicateurs axés sur l'expérience* : Ces indicateurs mesurent les expériences réelles de corruption des citoyens ou des entreprises, soit pour avoir offert des dessous de table, soit pour en avoir reçu.
- *Les indicateurs indirects (proxy indicators)* : Sur la base de la croyance que la corruption est impossible à mesurer empiriquement, les indicateurs indirects évaluent la corruption indirectement en regroupant de nombreuses « voix » et signaux de corruption ou en mesurant son opposé : anti-corruption, bonne gouvernance et mécanismes de responsabilité publique.
- *Les indicateurs sensibles à la pauvreté (pro-poor) et à l'égalité homme-femme (gender-sensitive)* : Un indicateur sensible à la pauvreté exige de se focaliser sur ceux qui vivent dans la pauvreté, et un indicateur sensible au genre considère les expériences et intérêts qui diffèrent entre femmes et hommes. De tels indicateurs sont utiles pour suivre les impacts potentiellement différents que les mécanismes et les processus de gouvernance ont sur des groupes sociaux différents.

La terminologie employée autour des différents types d'indicateurs reste subjective, mais l'essentiel est de disposer d'un ensemble varié d'indicateurs plutôt que de ne s'appuyer que sur un seul type d'indicateurs, par exemple ceux qui se fondent sur la perception. Malheureusement, en matière de lutte contre la corruption, les indicateurs purement empiriques et « objectifs », qui se basent sur des faits tangibles, sont rares. Pour cette raison, les sciences sociales ont donc largement recours aux indicateurs indirects qui, utilisés correctement, peuvent donner de très bonnes estimations de la réalité. L'interprétation d'un indicateur doit toutefois tenir compte de la nature de ce dernier. L'indicateur du « nombre d'affaires de corruption traduites en justice », par exemple, ne saurait constituer un indice du niveau de corruption dans le pays, étant donné qu'une hausse du nombre d'affaires traitées par les tribunaux peut être le signe d'une augmentation de la corruption, d'un regain de confiance dans les instances judiciaires, ou les deux à la fois ; on peut donc plutôt parler ici d'indice des efforts réalisés

par les AAC et la justice. De la même façon, rappelons que les indicateurs fondés sur la perception traduisent l'opinion subjective d'individus. Tout impact réel sur la corruption met un certain temps à se refléter dans l'opinion du public, c'est pourquoi les indicateurs de perception ne révéleront aucune corrélation directe (Galtung, 2005). En somme, le recoupement de différents types d'indicateurs contribue à la validité et à la fiabilité des conclusions d'une évaluation.

Le catalogue se concentre sur l'élaboration d'indicateurs, mais fournit également quelques appréciations sur la collecte de données, dans la colonne « sources de vérification ». Il existe plusieurs façons de mesurer l'impact et de suivre les progrès accomplis vis-à-vis des indicateurs sélectionnés. Parmi les méthodes les plus fréquentes figurent les sondages d'opinion publique, les diagnostics du secteur public, les enquêtes auprès du secteur privé, les enquêtes combinées, les enquêtes transnationales, les enquêtes sensibles à la pauvreté et à l'égalité homme-femme et enfin, les enquêtes sectorielles (comme les Enquêtes de suivi des dépenses publiques).²⁶ Si la collecte des données intègre également des questions sur la corruption, les statistiques nationales peuvent être utiles et permettre d'effectuer une comparaison des statistiques de différentes sources, comme c'est le cas des « statistiques miroir » en matière d'échanges commerciaux. Enfin, les AAC peuvent également collecter leurs propres données, que nous désignons tout simplement comme les « archives internes » dans le catalogue. Citons par exemple « l'enquête indépendante entreprise chaque année par l'ICAC à Hong-Kong, qui mesure, entre autres, le niveau de confiance entre l'ICAC et le public, le taux de poursuites engagées et le niveau, les types, les foyers et les causes de la corruption ». (ONU DC, 2002, 257)

7.2 Exemples d'indicateurs employés par trois AAC

Les trois AAC mentionnées dans ce chapitre sont représentatives de types d'organisations très disparates. Le Bureau britannique des fraudes graves (SFO) est spécialisé dans la criminalité financière et n'a jamais bénéficié de soutien de la part de donateurs. Le Bureau de prévention et de lutte contre la corruption (KNAB), en Lettonie, a pendant un temps reçu une assistance technique financée par des donateurs, mais s'est peu à peu consolidé et fournit aujourd'hui des services d'expertise à d'autres AAC. Enfin, certains indicateurs de réussite de la Commission indépendante contre la corruption, à Hong-Kong, sont issus de son rapport annuel et de ses critères de performance.

Les indicateurs que nous allons présenter ici illustrent comment un petit nombre d'AAC mesurent leurs performances. Ils ne constituent qu'une sélection sur l'ensemble des indicateurs que ces agences utilisent, et n'ont pas été choisis sur la base des critères SMART ou CREAM (voir chapitre 4.2.1). Le catalogue d'indicateurs présenté en annexe 1, en revanche, s'inspire d'un grand nombre d'expériences internationales, et notamment du travail de l'International Corruption Hunters Alliance, qui a permis de rassembler les indicateurs de performance actuellement utilisés par les AAC partout dans le monde. Le catalogue se limite en outre aux indicateurs conçus selon les critères mentionnés plus haut.

7.2.1 Le Bureau des fautes graves, au Royaume-Uni

Bien que le SFO, au Royaume-Uni, intervienne essentiellement sur la question de la criminalité financière et ait donc une mission plus ciblée que la plupart des AAC des pays développés et en transition, la corruption représente cependant une part importante de ses attributions. Comptant 307

²⁶ TI et le PNUD proposent une liste de ces différents outils (2007, 7–11).

employés permanents, le SFO mesure ses performances par le biais, entre autres, des indicateurs suivants :²⁷

- Montant de l'allocation budgétaire
- Nombre total d'employés
- Nombre d'affaires en cours
- Prix par habitant (utilisé comme mesure du rapport coût-bénéfice)
- Nombre de procès
- Nombre de condamnations
- Montant des indemnités versées aux victimes (compensations financières)
- Montant des sommes recouvrées par le biais des amendes imposées aux entreprises et réinjectées dans le système judiciaire (recouvrement pénal des fonds)
- Montant des sommes recouvrées dans des affaires par le biais de moyens civils (recouvrement civil des fonds)
- Montant des sommes saisies dans le cadre d'enquêtes financières
- Nombre de demandes d'informations traitées provenant d'autres pays au sujet d'affaires de fraude et de corruption (coopération internationale)

7.2.2 Le Bureau de prévention et de lutte contre la corruption, en Lettonie

Fondé en 2002, le KNAB, en Lettonie, compte aujourd'hui 142 employés. Les indicateurs de performance utilisés par le Bureau sont disponibles dans son rapport narratif sur les *Progrès et résultats en matière de prévention et de lutte contre la corruption en Lettonie* (KNAB, 2009). Bien que les indicateurs de performance n'y soient pas formulés de manière explicite, la liste suivante peut en être déduite :

- Nombre de propositions législatives en matière de lutte contre la corruption
- Nombre de projets de réglementation présentés
- Nombre de nouvelles lois déposées
- Nombre de révisions de lois adoptées par le Parlement
- Montant des financements illégaux de partis politiques découverts
- Montant des financements illégaux de partis politiques remboursés par les partis
- Nombre de déclarations de patrimoine analysées
- Sommes gagnées par des individus ayant enfreint les lois sur les conflits d'intérêts et les emplois supplémentaires et ayant été repérés par l'AAC
- Sommes recouvrées par l'AAC
- Nombre d'individus condamnés suite aux enquêtes menées par l'AAC

²⁷ Consulter le site Internet du SFO sur <http://www.sfo.gov.uk/our-work/our-performance.aspx>, qui propose le document « Achievements in 2009–10 ». Consulter également le rapport annuel 2009–2010 de SFO au Parlement, p. 23–25, sur <http://www.sfo.gov.uk/media/112684/sfo%20annual%20report%202009-2010.pdf>.

- Nombre de procédures pénales transmises au bureau du procureur
- Nombre de fonctionnaires formés sur les questions de conflits d'intérêts et d'éthique et sur les mesures anti-corruption internes
- Nombre de demandes internationales reçues par l'AAC pour fournir ses services d'expertise à l'étranger
- Nombre de délégations d'AAC étrangères reçues

Le KNAB et le STT lituanien (Service d'enquêtes spéciales), tous deux des AAC, utilisent les indicateurs suivants pour mesurer leurs performances comparatives dans le rapport *Normes communes et bonnes pratiques pour les agences anti-corruption* produit par le groupe de travail des Partenaires européens contre la corruption (EPAC, 2008) :

- Nombre d'enquêtes ouvertes
- Nombre d'affaires déferées devant le parquet

7.2.3 La Commission indépendante contre la corruption, à Hong-Kong

Les critères de performance peuvent également faire office d'indicateurs. L'ICAC de Hong-Kong, qui compte près de 1200 employés, s'appuie sur les critères de performance suivants qui, en fonction du nombre de fois où ils ont été atteints ou non, peuvent être traduits en indicateurs :²⁸

- Traiter les signalements d'actes de corruption dans les 48 heures (indicateur : ratio du nombre de signalements d'actes de corruption traités dans les 48 heures par rapport au nombre de signalements d'actes de corruption non traités dans les 48 heures)
- Traiter les signalements ne concernant pas des affaires de corruption dans un délai de deux jours ouvrables (indicateur : ratio du nombre de signalements ne concernant pas des affaires de corruption traités dans les deux jours ouvrables par rapport au nombre de signalements ne concernant pas des affaires de corruption non traités dans les deux jours ouvrables)
- Traiter les demandes de conseils en matière de prévention de la corruption dans un délai de deux jours ouvrables (indicateur : ratio du nombre de demandes de conseils traitées dans les deux jours ouvrables par rapport au nombre de demandes de conseils non traitées dans les deux jours ouvrables)
- Traiter les demandes d'informations ou de services pédagogiques en matière de lutte contre la corruption dans un délai de deux jours ouvrables (indicateur : ratio du nombre de demandes de services traitées dans les deux jours ouvrables par rapport au nombre de demandes de services non traitées dans les deux jours ouvrables)

Le rapport annuel de l'ICAC apporte d'autres exemples d'indicateurs qui ne sont pas liés aux services fournis mais plutôt axés sur les résultats obtenus (ICAC, 2009) :

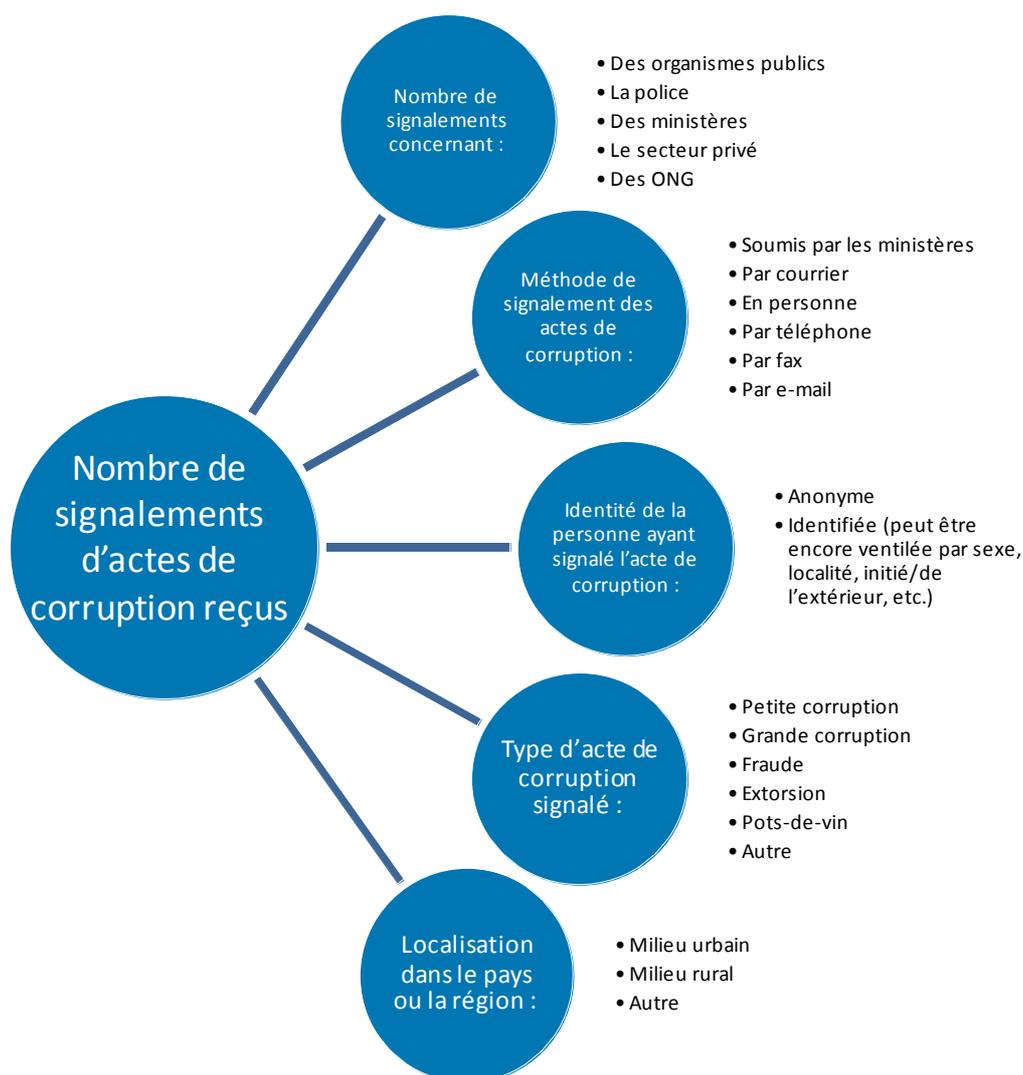
- Taux de condamnation basé sur les affaires traitées
- Nombre de visiteurs pendant la Journée portes ouvertes
- Nombre de signalements d'actes de corruption reçus (ventilés)

²⁸ Consulter le site Internet de l'ICAC sur les critères de performance, http://www.icac.org.hk/en/about_icac/mp/index.html. D'autres critères de performance sont disponibles dans ICAC (2011). http://www.icac.org.hk/filemanager/en/Content_1017/pledge_eng.pdf

- Nombre de signalements pouvant donner lieu à des suites
- Nombre d'enquêtes en cours
- Nombre d'enquêtes ouvertes et menées à terme
- Nombre total de dossiers
- Nombre de poursuites engagées (ventilées en fonction du type de corruption)
- Pourcentage de la population se disant confiante vis-à-vis de l'ICAC
- Nombre de visiteurs étrangers à l'ICAC

Le rapport annuel de l'ICAC suggère d'autres indicateurs, par exemple concernant les formations et le développement des compétences ou encore, la prévention et l'éducation à la corruption (étudiants formés, nombre d'universités intégrant la question de l'éthique à leurs cursus, etc.).

D'après l'ICAC, il est préférable de ventiler les indicateurs, de manière à mettre en valeur les principales différences entre les types de corruption, et donc de mieux appréhender le problème et d'optimiser la lutte contre la corruption. Le schéma 5 montre un exemple de ventilation de l'indicateur « nombre de signalements d'actes de corruption reçus ».

Schéma 5. Ventilation d'un exemple d'indicateur

7.3 Recommandations pour l'utilisation du catalogue d'indicateurs

Avant de sélectionner des indicateurs dans le catalogue, il est important de s'assurer qu'ils correspondent bien aux besoins de l'AAC concernée et aux impacts/objectifs, réalisations et produits visés – lesquels doivent être clairement définis au préalable.

Certaines AAC poursuivent une mission spécifique, qui appelle l'élaboration d'indicateurs uniques. Par exemple, les agences qui n'exercent que des fonctions de conseil stratégique et de coordination du système d'intégrité national seraient dans l'impossibilité d'utiliser des indicateurs conçus à des fins de prévention et de répression. Une AAC peut également se doter d'un ensemble d'indicateurs ciblant certaines de ses faiblesses, perçues ou réelles. Par exemple, une agence très productive au niveau administratif mais limitée en matière de répression des infractions pourra considérer le taux de condamnations atteint comme un indicateur clé de performance. Si les activités de l'agence sont trop centrées sur la petite corruption et ne permettent pas l'arrestation des grands délinquants, le pourcentage d'affaires impliquant des biens de grande valeur et/ou permettant la saisie de ces derniers peut alors constituer un indicateur de performance (facteur stimulant qui marquerait la volonté de lutter contre la grande corruption).

Dans l'annexe 1, les indicateurs sont classés d'après les fonctions de base que les AAC exercent habituellement :

- *La gestion globale* : se rapporte aux performances organisationnelles et aux pratiques de gestion telles que la stratégie organisationnelle, les procédures d'exploitation courantes, les mécanismes de contrôle internes, les procédures de signalement et la standardisation des méthodes de travail par la diffusion de documentation et de lignes directrices (manuels, guides), etc.
- *La production, la gestion et le partage des connaissances* : le système de S&E est une composante essentielle de cette fonction, mais celle-ci recouvre également les procédures d'analyse, la gestion et les pratiques de diffusion de tous les autres services.
- *L'application de la loi* : cette fonction englobe les trois dimensions clés de la répression que sont la collecte et l'analyse de données, l'ouverture d'enquêtes et l'engagement de poursuites.
- *La législation* : inclut l'analyse et le remaniement des actes législatifs en matière de lutte contre la corruption, les vérifications des critères de référence / de la conformité d'après les normes internationales, et l'analyse des défauts de mise en œuvre de la législation nationale.
- *La prévention* : renvoie non pas à l'application de la loi, mais à la prévention de cette dernière, qui passe notamment par l'éducation civique ainsi que par différentes mesures de promotion de l'intégrité comme la mise en place de codes de conduite, de régimes d'intégrité publique, de déclarations de patrimoine, etc.
- *La coopération entre agences* : les AAC font souvent office de points de convergence pour différentes institutions pénales, judiciaires et d'autres secteurs publics impliquées dans la lutte contre la corruption. La promotion de la coopération entre agences vise généralement à améliorer les procédures et les pratiques de traitement des dossiers.
- *La coopération internationale* : cette fonction recouvre les procédures internationales en matière de lutte contre la corruption transnationale, d'échanges de renseignements et de promotion de l'apprentissage.
- *La coopération avec la société civile* : méthodes de travail impliquant des acteurs de la société civile n'entrant pas dans les catégories susmentionnées.
- *La coopération avec le secteur privé* : méthodes de travail impliquant les associations professionnelles ou les sociétés individuelles n'entrant pas dans les catégories susmentionnées.

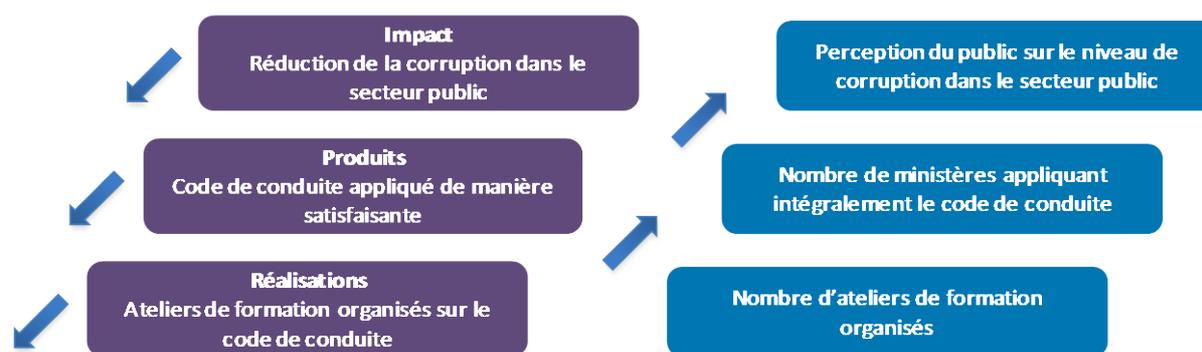
La majeure partie des objectifs, réalisations ou produits se rapportent à ces fonctions générales. Une AAC visant un résultat particulier en matière de prévention de la corruption pourra donc concentrer ses recherches dans le tableau fourni en annexe 1 en ciblant uniquement les indicateurs relatifs à cette fonction. Chaque indicateur est par ailleurs rattaché à une sous-fonction spécifique, de manière à clarifier les différents aspects de la fonction principale que l'indicateur doit mesurer. A titre d'indication pour la collecte et l'analyse des données, pour chaque indicateur, le tableau fait apparaître des sources de vérification potentielles. En dernier lieu, la grille propose différents commentaires et postulats.

Le catalogue d'indicateurs ne représente pas une liste imposée, mais plutôt un éventail de possibilités. Seuls les indicateurs qui reflètent le mieux les fonctions et les priorités de l'AAC évaluée doivent être sélectionnés. Le tableau contient peu d'indicateurs qualitatifs, non pas parce qu'ils ont moins de valeur que les indicateurs quantitatifs, mais plutôt parce qu'ils tendent généralement à mesurer des

réalisations fortement dépendantes du contexte, moins faciles à intégrer à un catalogue général, que l'on décompose généralement par étapes.²⁹

Le catalogue que nous présentons en annexe 1 est construit selon la méthode d'un modèle de cadre logique normal, basé sur une théorie du changement. Ainsi, les produits mènent à des réalisations, lesquelles entraînent des impacts. La construction de cadres logiques et de séries d'indicateurs exige d'établir une logique interne connectant les niveaux des produits, des réalisations et des impacts/objectifs. Pour la théorie du changement, le processus se fait généralement en commençant par la fin, c'est-à-dire qu'à partir des impacts visés, sont d'abord définis les réalisations nécessaires, puis les produits requis pour produire ces réalisations. L'élaboration des indicateurs doit en revanche suivre la démarche inverse et commencer par la base, en définissant les indicateurs de produits, puis de réalisations et enfin, d'impacts. C'est bien sûr une vision schématisée, dans la mesure où une logique interne solide se doit d'être élaborée à la fois en remontant et en descendant la chaîne de causalité, mais cela peut constituer un principe directeur utile. Le schéma 6 se base sur l'exemple des codes de conduite, mais le modèle peut être appliqué à tout type d'intervention.

Schéma 6. Le modèle logique connecté aux indicateurs



Source : Adapté d'une diapositive présentée lors de la formation "Monitoring & Evaluation for Results", organisée par le Groupe d'évaluation de l'Institut de la Banque mondiale, Washington, DC, 2009.

A partir des principes que nous venons de voir, le schéma 7 explique comment le catalogue fonctionne. Comme le montrent les flèches violettes, l'impact visé détermine les réalisations nécessaires et les produits à réaliser, et donc les fonctions que les AAC doivent exercer. Les flèches bleues indiquent que le choix / l'élaboration des indicateurs doit suivre une logique ascendante. Les flèches rouges décrivent l'ordre normal de l'élaboration des indicateurs : on établit d'abord l'impact visé (au niveau des objectifs) et les fonctions (au niveau des réalisations et des produits) à mesurer ; ensuite, on définit les indicateurs, on détermine les sources de vérification et on éclaircit les postulats. Les niveaux des réalisations et des produits comportent une étape supplémentaire, du fait que les fonctions soient tout d'abord traduites en réalisations/produits souhaités puis concrétisées en indicateurs. Pour bien utiliser ce catalogue, il est souhaitable de classer les indicateurs sous les fonctions générales que les AAC exercent normalement. Cela n'est toutefois pas nécessaire au niveau

²⁹ Citons par exemple l'adoption d'une législation sur les conflits d'intérêts au parlement. Etape 1 : l'AAC soumet un projet de loi en matière de conflits d'intérêts au parlement. Etape 2 : adoption de la loi sur les conflits d'intérêts par le parlement. Etape 3 : élaboration des arrêtés et des règlements sur la base de la loi adoptée. Etape 4 : Mise en application de la loi sur les conflits d'intérêts dans le service public. Le cadre logique du DFID est particulièrement utile à la mise en pratique d'indicateurs qualitatifs (DFID, 2009). <http://www.dfid.gov.uk/Documents/publications/1/how-to-guid-rev-log-fmwk.pdf>

des impacts/objectifs, où l'usage est d'utiliser les instruments standards de mesure de la corruption. Notons que le niveau des objectifs est généralement considéré comme extérieur à la « sphère d'influence » de l'intervention, c'est-à-dire qu'on n'en attend aucune attribution directe.

Schéma 7. Comment utiliser le catalogue d'indicateurs fourni en annexe 1

AU NIVEAU DES IMPACTS / DES OBJECTIFS				
Impacts recherchés		Indicateurs possibles	Sources de vérification	Commentaires, postulats
Impact 1		Indicateur 1a Indicateur 1b	Source 1a Source 1b	
AU NIVEAU DES REALISATIONS				
Fonctions	Résultats visés	Indicateurs possibles	Sources de vérification	Commentaires, postulats
Fonction 1	Résultat 1.1	Indicateur 1.1a	Source 1.1a	
	Résultat 1.2	Indicateur 1.1b Indicateur 1.2a Indicateur 1.2b	Source 1.1b Source 1.2a Source 1.2b	
Fonction 2	Résultat 2.1	Indicateur 2.1a	Source 2.1a	
	Résultat 2.2	Indicateur 2.1b Indicateur 2.2a Indicateur 2.2b	Source 2.1b Source 2.2a Source 2.2b	
AU NIVEAU DES PRODUITS				
Fonctions	Résultats visés	Indicateurs possibles	Sources de vérification	Commentaires, postulats
Fonction 1	Accomplissement 1.1.1	Indicateur 1.1.1a	Source 1.1.1a	
	Accomplissement 1.1.2	Indicateur 1.1.1b	Source 1.1.1b	
	Accomplissement 1.2.1	Indicateur 1.1.2a	Source 1.1.2a	
	Accomplissement 1.2.2	Indicateur 1.1.2b Etc.	Source 1.1.2b Etc.	
Fonction 2	Accomplissement 2.1.1	Indicateur 2.1.1a	Source 2.1.1a	
	Accomplissement 2.1.2	Indicateur 2.1.1b	Source 2.1.1b	
	Accomplissement 2.2.1	Indicateur 2.2.1a	Source 2.2.1a	
	Accomplissement 2.2.2	Indicateur 2.2.2b Etc.	Source 2.2.2b Etc.	

L'application de la logique décrite ci-dessus lors de l'utilisation du catalogue d'indicateurs garantira la définition de séries d'indicateurs cohérentes. Ces dernières doivent couvrir les niveaux des impacts, des réalisations et des produits, et être adaptées aux fonctions visées par l'AAC.

8. Capitaliser les connaissances acquises dans le cadre du S&E

Au cours des vingt dernières années, les AAC se sont multipliées. On en compte aujourd'hui une centaine, réparties sur tous les continents. A ce nombre, il faut ajouter la multitude d'organismes territoriaux qui n'entrent pas dans les estimations chiffrées, dont le caractère est très disparate d'un pays à un autre. Certains Etats dotent leurs agences de pouvoirs d'investigation et même de la capacité à engager des poursuites, alors que d'autres leur confèrent plutôt un rôle préventif, pédagogique, politique ou de coordination. Certaines de ces agences sont récentes, quand d'autres ont déjà une longue expérience. Nombre d'évaluations de ces AAC ont été effectuées, mais qu'en a-t-il été retenu exactement ?

La communauté internationale s'est activement employée à populariser les AAC comme étant la solution institutionnelle au problème de la corruption. L'OCDE (1996, 14) fut la première qui suggéra aux Etats membres de créer ce type de service indépendant et spécialisé comme partie intégrante de leurs « infrastructures éthiques ». Par la suite, d'autres initiatives ont vu le jour au niveau régional et international.³⁰

Cependant, même aujourd'hui, les débats autour des AAC tendent à répondre à une logique binaire : certains les perçoivent comme étant la solution, d'autres pas. La question n'est pourtant pas de savoir *si* les AAC sont de bonnes ou de mauvaises institutions, mais plutôt *pourquoi, dans quelle mesure, dans quels contextes*, etc. Pour éclairer le débat, les évaluations d'AAC ne devraient pas seulement constituer des audits d'organisations reposant sur l'usage de listes de contrôle ou sur des impératifs de redevabilité pour évaluer si les agences disposent bien des pouvoirs et des structures institutionnelles qui leur ont été conférés. Au lieu de cela, elles devraient chercher à définir les déterminants de bonnes interventions anti-corruption, à clarifier l'impact réel des AAC et à étudier les moyens d'améliorer leur travail. En somme, elles devraient s'inscrire dans une démarche de capitalisation des connaissances. A cet effet, la participation de toutes les parties prenantes et l'instauration de services de S&E sont indispensables (voir chapitres 5.5 et 4.2). Tout processus de capitalisation des connaissances acquises dans le cadre du S&E devrait inclure les quatre composantes suivantes :

- *Des mécanismes intégrés de gestion de la remontée et de diffusion des informations* : Une fois terminées, les évaluations ne devraient être ni simplement archivées ni diffusées à un noyau restreint de destinataires. Elles devraient relever d'une procédure prévoyant des mécanismes de diffusion des rapports et de traitement des commentaires et opinions formulés à leur sujet. Des évaluations bien construites contiennent des informations importantes pour les programmes et les stratégies à venir. Ici encore, la participation est vitale pour garantir la transparence du processus. Les mécanismes formels de gestion de la remontée de l'information pourraient prévoir des procédures de révision, des révisions par les pairs, des séminaires et ateliers, etc. Les AAC et les donateurs devraient également encourager les activités informelles telles que la mise en réseaux et la communication interne, vecteurs de diffusion des idées et des informations. Comme le souligne l'OCDE, « pour être efficaces, les mécanismes de gestion du feedback doivent reposer sur des ressources humaines et budgétaires et bénéficier du soutien des dirigeants et des autres

³⁰ La Convention pénale du Conseil de l'Europe sur la corruption, la Convention interaméricaine contre la corruption, le Protocole anti-corruption de la Communauté de développement de l'Afrique australe, la Convention de l'Union africaine sur la prévention et la lutte contre la corruption et enfin, la Convention des Nations unies contre la corruption (CNUCC) font tous référence à la nécessité d'instaurer des autorités indépendantes de lutte contre la corruption.

acteurs impliqués » (2010a, 10). Ces mécanismes devraient donc être planifiés et budgétisés au titre de la procédure d'évaluation, et non pas pensés après coup. Les AAC et les membres des organismes donateurs chargés d'intégrer ces mécanismes de gestion du feedback aux évaluations peuvent consulter la publication de l'OCDE (2001) *Evaluation Feedback for Effective Learning and Accountability*.

- *Le S&E en tant qu'instrument de capitalisation* : La principale difficulté reste sans doute d'instaurer une culture au sein des AAC qui intègre le S&E comme un outil d'apprentissage utile plutôt que comme une contrainte malvenue. Cela passe d'abord par la promotion de la participation : tous les membres des AAC devraient pouvoir intervenir au cours des évaluations et signaler d'éventuels problèmes. Comme nous allons le voir ci-dessous, il convient ensuite de veiller à ce que les évaluations puissent alimenter le travail quotidien du personnel.
- *Le S&E doit être utile* : Il peut sembler superflu de souligner le fait que les systèmes de suivi et d'évaluation doivent être utiles. Cependant, une évaluation dont l'objectif et les destinataires ne sont pas définis est d'une utilité discutable. Une bonne évaluation est conçue, réalisée et divulguée dans un but précis et répond aux besoins de ses destinataires. Pour que les conclusions, les recommandations et les leçons d'une évaluation puissent être capitalisées, elles doivent être à la fois claires, pertinentes, ciblées et réalisables (OCDE, 2010a, 29).
- *Tirer des leçons des recommandations* : Les AAC doivent veiller à ce que les personnes responsables répondent systématiquement aux recommandations des évaluateurs. Si la direction d'une AAC n'est pas d'accord avec les recommandations reçues, elle doit le faire savoir au cours des procédures de feedback ou de révision. Les AAC devraient développer un système formel de réponse et de suivi (par exemple, par le biais d'un plan d'action) afin de systématiser la mise en application des recommandations reçues. Le service de S&E et la direction devraient veiller conjointement à ce que toutes les actions de suivi convenues soient bien entreprises, de manière à pouvoir rendre compte de leur mise en œuvre. Des évaluations devraient ensuite être effectuées pour contrôler la réaction de la direction aux recommandations, les éventuelles mesures mises en place et les effets obtenus. Dans le cas où ces derniers ne correspondent pas aux effets désirés, ils devront être justifiés.

La capitalisation des connaissances est sans doute l'aspect le plus négligé du S&E. Si les AAC doivent répondre à des impératifs de redevabilité, ces derniers ne doivent pas éclipser la nécessité d'assimiler les méthodes qui fonctionnent et les raisons de leur succès.

Annexe 1. Catalogue d'indicateurs

AU NIVEAU DES IMPACTS / DES OBJECTIFS			
Impact recherché	Indicateurs possibles	Sources de vérification	Commentaires, postulats
Réduction des délits de corruption et des autres infractions qui y sont liées	<p>Nombre de délits de corruption signalés</p> <p>Nombre de condamnations</p> <p>Nombre de mesures disciplinaires</p> <p>Taux de représailles</p>	<ul style="list-style-type: none"> ▪ Statistiques nationales ▪ Rapports de police ▪ Procès-verbaux judiciaires ▪ Baromètre mondial de la corruption de Transparency International 	<p>Peut être encore divisé par type de délits conformément à la définition juridique qui en est faite ou par type de corruption. Lors de l'examen de procès-verbaux judiciaires, clarifier le rôle des AAC par une analyse du contenu.</p> <p>Le suivi de cet indicateur doit être effectué sur une longue période. On peut attendre des AAC qu'elles contribuent à un nombre plus élevé de procès, mais de nombreux autres intervenants entrent en jeu qui en déterminent l'issue.</p>
Amélioration de la transparence et diminution de la corruption dans les services publics	<p>Pourcentage de la population locale satisfaite des services publics en termes de transparence et de corruption</p>	<ul style="list-style-type: none"> ▪ Enquêtes sur les services fournis ▪ Enquête de satisfaction des usagers ▪ Avis d'experts 	<p>La formulation du questionnaire sera adaptée en fonction de la situation et des institutions locales.</p> <p>Il est important d'assurer la cohérence de l'évaluation en effectuant le suivi des mêmes institutions au fil du temps. Centrer le travail des AAC sur les institutions ciblées.</p>
Réduction du niveau de corruption perçue	<p>Perception du niveau de corruption ressortant de sondages d'opinion</p>	<ul style="list-style-type: none"> ▪ Sondages d'opinion ▪ Baromètres régionaux (Afrique, Asie, Amérique latine, Europe) ▪ IPC, Indice de corruption des pays exportateurs et Baromètre mondial de la corruption de Transparency International ▪ Rapports d'ONG ▪ Veille médiatique 	<p>Mesure de la perception de la corruption du grand public</p>

AU NIVEAU DES IMPACTS / DES OBJECTIFS

<p>Assainissement des activités économiques (diminution de la corruption)</p>	<p>Montant des pots-de-vin versés par les entreprises</p>	<ul style="list-style-type: none"> ▪ Indice de compétitivité mondiale (Global competitiveness Index) ▪ Etude mondiale de PriceWaterhouseCoopers sur la fraude en entreprise ▪ Enquête de la Banque mondiale sur le cadre de l'activité économique et la performance des entreprises (BEEPS) ▪ Enquête de la Banque mondiale et de la Société financière internationale (groupe de la Banque mondiale) auprès des entreprises ▪ Enquête sur le climat des investissements du Département d'Etat nord-américain (US State Department Investment Climate Survey) 	<p>Pour cet indicateur, le postulat de départ est qu'un milieu des affaires moins corrompu facilite la croissance économique.</p>
<p>Instauration d'un climat de confiance entre les individus pour une diminution de la corruption</p>	<p>Niveau de confiance des individus entre eux</p>	<ul style="list-style-type: none"> ▪ Données du sondage du Quality of Government Institute-Expert Survey dataset 	<p>Il existe différentes mesures de la confiance entre les individus, qui restent à préciser. Peut également être associée avec la confiance vis-à-vis des institutions gouvernementales, au niveau des résultats.</p>
<p>Sensibilisation du public aux effets négatifs de la corruption</p>	<p>Niveau de conscience des conséquences négatives de la corruption, d'après des sondages d'opinion de la population</p>	<ul style="list-style-type: none"> ▪ Enquêtes sur le niveau de sensibilisation du public 	<p>Cet indicateur vise surtout à évaluer le degré de connaissance des conséquences négatives de la corruption auprès de la population.</p>

AU NIVEAU DES REALISATIONS

Fonction	Réalisations visées	Indicateurs possibles	Sources de vérification	Commentaires, postulats
La gestion globale	Les donateurs considèrent que l'AAC est performante	Rapports de donateurs reconnaissant la contribution positive de l'AAC (oui/non)	<ul style="list-style-type: none"> ▪ Etudes diagnostiques de la Banque mondiale pour la bonne gouvernance et la lutte contre la corruption (gouvernance et anti-corruption, GAC) ▪ Rapports du Groupe d'Etats contre la corruption (GRECO) ▪ Evaluations externes 	Au niveau des résultats généraux, les performances des AAC peuvent être évaluées par des spécialistes, généralement mandatés par les organismes donateurs.
	La population considère que l'AAC est performante	Pourcentage de la population se disant satisfaite des performances de l'AAC	<ul style="list-style-type: none"> ▪ Enquêtes 	La satisfaction du public vis-à-vis de l'AAC est importante, mais il faut rester conscient du fait qu'il faut des années avant que la population n'acquière une réelle connaissance des activités des AAC.
	L'AAC a la confiance du public	Niveau de confiance du public vis-à-vis de la mission de lutte contre la corruption de l'AAC	<ul style="list-style-type: none"> ▪ Enquêtes 	Sous-ensemble de l'indicateur précédent. Le niveau de confiance entre le public et les AAC étant déterminant dans le succès de la lutte contre la corruption, il devrait être suivi de près. Pourrait être détaillé de manière à évaluer des mesures spécifiques d'intégrité, comme la confiance vis-à-vis du mandataire.

AU NIVEAU DES REALISATIONS

Les médias considèrent que l'AAC est compétente et qu'elle obtient de bons résultats	Nombre d'articles de presse positifs et négatifs sur les AAC, que ce soit sur des affaires particulières ou de l'information générale	<ul style="list-style-type: none"> ▪ Veille médiatique, analyse des articles de presse 	Dans les pays où les médias sont relativement indépendants, la performance d'une AAC peut se mesurer au traitement que ces derniers lui réservent. Cet indicateur peut être ventilé en fonction de l'exposition médiatique des affaires traitées. Comparer leurs succès et leurs échecs permet d'enrichir les informations analysées.
La structure, les systèmes et les procédures institutionnelles de l'AAC fonctionnent bien	<p>Nombre de défaillances révélées dans des rapports d'évaluation d'experts</p> <p>Nombre de défaillances identifiées comme graves et appelant une action immédiate</p>	<ul style="list-style-type: none"> ▪ Evaluations des rapports d'experts 	Les évaluations pourraient vérifier que les fonctions de chaque poste fassent l'objet d'une description claire et soient intégrées à un organigramme, réviser les descriptions de postes, les procédures de recrutement et de promotion, etc.
L'AAC emploie du personnel compétent	Niveau de compétence du personnel en matière de corruption	<ul style="list-style-type: none"> ▪ Comparaison des niveaux de compétences individuelles par le biais de tests avant/après 	Cet indicateur dépend du genre de formations proposées.
L'AAC veille à la formation de son personnel	Niveau de satisfaction du personnel vis-à-vis des formations proposées, des programmes et du matériel pédagogique et des modules de formation	<ul style="list-style-type: none"> ▪ Questionnaires de satisfaction sur les formations effectuées 	La formulation exacte de cet indicateur dépend des modalités de formation de l'AAC concernée. Si les formations sont assurées par des membres internes de l'agence, le niveau de satisfaction vis-à-vis des formations proposées constitue alors un bon indicateur.

AU NIVEAU DES REALISATIONS

La production et la gestion des connaissances	L'AAC a de bonnes capacités de diagnostic et d'analyse	Nombre d'études effectuées ou mandatées par l'AAC	<ul style="list-style-type: none"> ▪ Informations internes 	Le nombre d'études est un indicateur sommaire de la qualité des capacités de diagnostic de l'AAC. Une évaluation plus approfondie serait ici nécessaire.
	L'AAC dispose d'un système opérationnel de S&E axé sur les résultats et a les capacités nécessaires pour l'utiliser	Système de S&E établi et opérationnel (oui/non)	<ul style="list-style-type: none"> ▪ Informations internes/évaluation d'expert externe 	L'emploi du terme « opérationnel » dans l'indicateur devrait au minimum être défini de manière à signifier que les indicateurs sont établis, les données systématiquement collectées, et une personne chargée de la collecte et du stockage systématiques des données.
	L'AAC dispose d'un système fonctionnel et sécurisé de gestion et de partage des connaissances	Système de gestion et de partage des connaissances établi et opérationnel, avec les dispositions nécessaires pour la protection des données (oui/non)	<ul style="list-style-type: none"> ▪ Informations internes/évaluation d'expert externe 	Cela demande d'évaluer les méthodes de partage des informations au sein de l'organisation ; les moyens d'accès du personnel aux informations non confidentielles ; si des réunions sont régulièrement organisées ; si la direction partage les informations en interne et avec le personnel, et de quelle manière, par voie de circulaires ou par d'autres moyens ; s'il existe des systèmes de bases de données et que les personnes chargées de les alimenter sont clairement identifiées ; qui peut y accéder, etc.

AU NIVEAU DES REALISATIONS

L'application de la loi	Appréciation quantitative et qualitative des poursuites pour des affaires de corruption	Nombre d'affaires de corruption signalées par l'AAC et transmises au bureau du procureur ou à un tribunal spécial	<ul style="list-style-type: none"> Statistiques des services de police, des tribunaux et sur les poursuites engagées 	Cet indicateur peut être ventilé afin d'obtenir des informations plus détaillées sur chaque étape de la procédure ou sur chaque type d'affaire. Par exemple, il peut être utile de connaître le type d'infractions transférées pour comprendre si l'AAC est efficace dans la répression de la grande corruption.	
	Meilleure intégration de la répression et de la prévention dans les procédures de travail de l'AAC	Nombre d'initiatives de la direction mises en œuvre et/ou d'activités de prévention entreprises sur la base des données transmises par la division chargée des activités de répression	<ul style="list-style-type: none"> Informations internes : révision des rapports annuels, des plans stratégiques, des plans d'action des différents services, etc. 	Cet indicateur vise à déterminer si les informations générées par les activités de répression (renseignement, enquêtes, poursuites) alimentent des réformes institutionnelles et procédurales spécifiques et/ou génériques, la définition des priorités et les décisions stratégiques, ou plutôt des activités de prévention et d'éducation.	
	Amélioration quantitative et qualitative des capacités d'enquête	Ratio des enquêtes menées par rapport aux poursuites engagées Ratios des enquêtes menées par rapport aux condamnations prononcées et des poursuites engagées par rapport aux condamnations prononcées		<ul style="list-style-type: none"> Informations internes Archives judiciaires 	Cet indicateur mesure non seulement l'efficacité des enquêteurs mais également celle du système de sélection des dossiers de l'AAC.
		Nombre d'affaires traitées faisant l'objet d'une enquête de l'AAC		<ul style="list-style-type: none"> Informations internes 	Cet indicateur mesure la charge de travail de l'AAC en termes d'enquêtes.

AU NIVEAU DES REALISATIONS

		Nombre d'affaires classées ou d'individus acquittés pour cause de vices de procédure ou d'enquête inefficace de l'AAC	<ul style="list-style-type: none"> Statistiques des services de police, des tribunaux et sur les poursuites engagées 	Le nombre d'erreurs procédurales graves ou de mauvaises pratiques d'investigation de l'AAC devrait être minimisé. Cet indicateur peut également refléter l'interaction entre le parquet/les tribunaux et l'AAC. En l'absence de coopération, l'indicateur ne sera pas bon.
		Nombre d'affaires classées ou d'individus acquittés en raison du manque de preuves	<ul style="list-style-type: none"> Statistiques des services de police, des tribunaux et sur les poursuites engagées 	Cet indicateur ne reflète pas nécessairement de mauvaises performances de l'AAC (il n'est pas toujours possible d'établir des preuves, par exemple si les accusés sont innocents) mais fournit des informations utiles.
	Niveau de condamnations élevé, y compris les condamnations de délits non mineurs	Nombre de condamnations ou pourcentage de condamnations sur le total des enquêtes menées	<ul style="list-style-type: none"> Archives judiciaires 	Cet indicateur ne reflète réellement les performances de l'AAC que si l'agence est investie de pouvoirs répressifs et que les tribunaux fonctionnent bien.
		Pourcentage de condamnations impliquant des biens de grande valeur et/ou leur saisie	<ul style="list-style-type: none"> Archives judiciaires 	Cet indicateur cible la corruption des élites et la grande corruption.
	Augmentation des saisies de biens	Valeur des biens saisis dans le cadre des enquêtes de l'AAC	<ul style="list-style-type: none"> Informations internes 	Plutôt que de se concentrer sur le nombre de saisies, la valeur totale fournit une mesure de l'impact global (100 petites saisies n'auront pas l'impact d'une grande saisie).

AU NIVEAU DES REALISATIONS

	Amélioration du système de recouvrement d'avoirs	<p>Nombre de défauts du système de recouvrement d'avoirs identifiés par des évaluations d'experts</p> <p>Nombre de défauts identifiés comme étant « graves » et appelant une réponse immédiate</p>	<ul style="list-style-type: none"> Documentation interne/évaluation d'expert externe 	<p>Les évaluations gagneraient à couvrir les lacunes typiques des systèmes de recouvrement d'avoirs : comment les avoires sont-ils évalués ; comment sont-ils gérés ; quelles possibilités la loi offre-t-elle de rendre des avoires liquides en raison d'une probable dépréciation ; une distinction claire est-elle faite entre la saisie et la confiscation ; les avoires sont-ils comptabilisés deux ou trois fois au cours des différentes étapes de la procédures (par exemple, la saisie, la confiscation, la cession) ; qu'advient-il des produits des délits après leur cession ; le public bénéficie-t-il de manière visible des biens recouvrés.</p>
	L'AAC contribue au recouvrement financier des avoires d'origine criminelle	Sommes recouvrées sous forme d'amendes et de compensations financières transférées par le biais du système judiciaire	<ul style="list-style-type: none"> Informations internes Archives judiciaires 	Cet indicateur mesure les avantages financiers que les avoires d'origine criminelle recouvrés revêtent pour l'Etat ou pour les individus.
La législation	Adoption de nouvelles lois nationales	Nombre de nouvelles propositions législatives en matière de lutte contre la corruption introduites dans la législation	<ul style="list-style-type: none"> Procès-verbaux officiels des séances parlementaires / journal parlementaire officiel Informations internes 	Cet indicateur peut également devenir plus qualitatif s'il souligne les domaines spécifiques dans lesquels de nouvelles législations sont nécessaires.
	Modification des lois nationales	Nombre d'amendements à la législation en vigueur adoptés par le parlement	<ul style="list-style-type: none"> Procès-verbaux officiels des séances parlementaires / journal parlementaire officiel 	Ici encore, plutôt que de comptabiliser le nombre d'amendements, on peut lister les amendements spécifiques que l'AAC souhaite cibler.

AU NIVEAU DES REALISATIONS

L'influence politique	L'AAC a contribué à la définition des priorités nationales	Nombre de contributions politiques apportées par l'AAC	<ul style="list-style-type: none"> ▪ Informations internes 	Les contributions politiques regroupent les propositions stratégiques, les lettres de plaidoyer, les publications dans les médias, etc.
	L'AAC a une influence politique positive	Nombre de documents stratégiques du gouvernement / d'actes législatifs sensiblement influencés par les contributions de l'AAC par an	<ul style="list-style-type: none"> ▪ Informations internes ▪ Procès-verbaux officiels ▪ Articles de presse 	« Sensiblement » implique qu'une contribution à un changement / une promulgation peut être attribuée à l'AAC par le biais de mesures objectives telles que les registres publics, la couverture dans les médias ou la traduction des variations politiques dans le plan stratégique ou de plaidoyer de l'AAC.
La prévention	L'AAC promeut un sens aigu du service public	Pourcentage des fonctionnaires conscients des dilemmes moraux et des directives à suivre en matière de corruption	<ul style="list-style-type: none"> ▪ Enquête ▪ Questionnaires sur les formations reçues 	Cet indicateur part du principe que l'AAC ou les autres institutions proposent des formations en matière d'éthique et de lutte contre la corruption.
		Pourcentage / nombre d'enquêtes approfondies sur des vices graves de procédures ou des manquements à l'éthique de hauts fonctionnaires entreprises intégralement et équitablement et menant à des conclusions transparentes et, si nécessaire, à la mise à exécution de la loi / à des peines dans les institutions ciblées par l'AAC	<ul style="list-style-type: none"> ▪ Archives ministérielles 	Si cet indicateur n'est pas déjà mesuré par les services publics, il représente une charge de travail conséquente pour l'AAC. Si la corruption du secteur public est une question importante pour l'agence, ce travail vaut la peine d'être entrepris, mais il doit entrer dans les priorités budgétaires de l'AAC.

AU NIVEAU DES REALISATIONS

L'AAC favorise la confiance vis-à-vis des services publics	Pourcentage d'usagers des services publics qui ont confiance dans les services publics	<ul style="list-style-type: none"> ▪ Enquête ▪ Enquête de satisfaction des usagers 	Cet indicateur se base sur la perception / l'expérience du public, il ne faut donc pas attendre qu'il traduise des changements rapides. Il peut aussi être uniquement appliqué aux institutions publiques avec lesquelles l'AAC travaille.
Réduction du temps / du coût réel d'obtention de licences ou de services	Temps / coût réel d'obtention d'une licence auprès de l'organisme chargé de sa délivrance ou connexion aux services par le biais d'une entreprise de services publics	<ul style="list-style-type: none"> ▪ Enquête 	Le coût réel inclut le prix officiel plus le montant moyen du pot-de-vin demandé.
L'AAC lutte contre le financement illégal des partis politiques	Montant des financements illégaux de partis politiques divulgués par l'AAC	<ul style="list-style-type: none"> ▪ Informations internes 	Si la lutte contre le financement illégal des partis politiques est une question prioritaire pour l'AAC, il est important de connaître le montant des fonds divulgués.
	Montant des financements illégaux de partis politiques remboursés par les partis	<ul style="list-style-type: none"> ▪ Informations internes ▪ Archives du ministère des Finances 	Le montant des financements remboursés par les partis politiques est révélateur à la fois des effets préventifs et répressifs du travail de l'AAC.
Réduction des problèmes de conflits d'intérêts dans l'administration publique	Sommes gagnées par des individus ayant enfreint les lois sur les conflits d'intérêts et les emplois supplémentaires et ayant été repérés par l'AAC	<ul style="list-style-type: none"> ▪ Informations internes 	Le nombre de délits de conflits d'intérêts pourrait constituer un indicateur alternatif.

AU NIVEAU DES REALISATIONS

	Les déclarations de patrimoine ont un effet dissuasif sur la corruption	Taux de déclarations de patrimoine remplies par les fonctionnaires	<ul style="list-style-type: none"> Informations internes 	Cet indicateur pourrait également prévoir de mesurer la qualité des déclarations effectuées. L'indicateur part du principe que l'AAC est responsable des déclarations de patrimoine ou les supervise. Des contrôles aléatoires de la conformité des déclarations de patrimoine pourraient permettre de mesurer la qualité du travail de l'AAC.
	Les audits de risques de corruption et/ou les régimes d'intégrité publiques contribuent au recul de la corruption	<p>Nombre/pourcentage d'institutions publiques ayant mis en œuvre la majorité des mesures recommandées par les audits de risques / les régimes d'intégrité publique dans les délais suggérés</p> <p>Perception du public / des fonctionnaires sur l'efficacité des audits de risques / des régimes d'intégrité</p>	<ul style="list-style-type: none"> Enquêtes 	Cet indicateur dépend entièrement du contexte et doit donc être adapté selon chaque cas.
La coopération entre agences	L'AAC assure la promotion de la coopération entre les institutions impliquées dans la lutte contre la corruption	La distribution des compétences parmi les institutions anti-corruption est claire, concertée et propice à la lutte contre la corruption (oui/non)	<ul style="list-style-type: none"> Evaluations externes Existence de manuels / directives / instructions entre les agences dans lesquels les compétences et les procédures de coopération sont spécifiées 	Cet indicateur qualitatif peut être approfondi davantage en détaillant les étapes représentant les activités requises pour parvenir à l'objectif final.

AU NIVEAU DES REALISATIONS				
	L'AAC coopère activement avec les autres institutions anti-corruption	Nombre d'initiatives / d'opérations / d'enquêtes conçues et entreprises conjointement	<ul style="list-style-type: none"> ▪ Informations internes 	Cet indicateur dépend bien sûr des pouvoirs d'investigation de l'AAC.
	Amélioration du traitement des dossiers	Nombre moyen de jours entre la constitution d'un dossier et sa prise en charge par les tribunaux Nombre de jours maximum garanti de traitement des dossiers	<ul style="list-style-type: none"> ▪ Informations internes ▪ Archives judiciaires ▪ Accords de coopération entre agences 	Ces indicateurs peuvent être adaptés de nombreuses manières en fonction des besoins. Par exemple, si certains ralentissements sont identifiés, les indicateurs peuvent se concentrer sur cette question.
La coopération internationale	Les informations font l'objet d'échanges internationaux fructueux	Nombre de demandes d'informations concernant des affaires de fraudes et de corruption adressées par d'autres pays et traitées par l'AAC	<ul style="list-style-type: none"> ▪ Informations internes 	Le fait que l'AAC reçoive des demandes d'informations d'institutions étrangères n'est qu'une supposition, mais elle est appropriée au niveau des résultats.
	Organisation et mise en œuvre d'opérations internationales communes	Nombre d'initiatives / d'opérations organisées et mises en œuvre conjointement avec des institutions anti-corruption étrangères	<ul style="list-style-type: none"> ▪ Informations internes 	Cet indicateur dépend bien sûr des pouvoirs d'investigation de l'AAC.
La coopération avec la société civile	La société civile a les capacités d'exiger du gouvernement qu'il réponde de la corruption	Nombre d'ONG compétentes dans l'analyse, le contrôle et la divulgation de la corruption au sein du gouvernement	<ul style="list-style-type: none"> ▪ Auto-évaluations des ONG 	Cet indicateur part du principe que l'AAC a formé le personnel de ces ONG.

AU NIVEAU DES REALISATIONS

	Les médias sont indépendants et libres	Nombre de journalistes assassinés, expulsés et/ou intimidés en raison de leur travail sur la corruption	<ul style="list-style-type: none"> ▪ Informations internes ▪ Statistiques de Reporters sans frontières 	Il existe de nombreux indices internationaux sur les médias qui disposent d'indicateurs et de résultats pour tout un ensemble de pays. Si le travail avec les médias fait partie des priorités de l'AAC, cette dernière peut également entreprendre de collecter ses propres informations.
		Monopole d'Etat à la radio et à la télévision (oui/non)	<ul style="list-style-type: none"> ▪ Informations internes ▪ Statistiques de Reporters sans frontières ▪ Statistiques de Freedom House 	
		Nombre de programmes / d'articles des médias censurés	<ul style="list-style-type: none"> ▪ Informations internes ▪ Statistiques de Reporters sans frontières 	
La coopération avec le secteur privé	L'AAC contribue au recul de la corruption dans le secteur privé	Nombre de signalements de corruption effectués par des entreprises et sur lesquels l'AAC enquête	<ul style="list-style-type: none"> ▪ Informations internes 	Il est important de ne mesurer que les questions que l'AAC peut influencer au niveau des résultats.
		Nombre de renvois effectués par l'AAC vers d'autres instances de justice pour des affaires de corruption du secteur privé	<ul style="list-style-type: none"> ▪ Informations internes 	Idéalement, il faudrait également mesurer le résultat final de ces renvois.
		Perception du secteur privé sur la contribution de l'AAC à la lutte contre la corruption	<ul style="list-style-type: none"> ▪ Enquêtes 	Cet indicateur peut être mesuré par le biais des enquêtes existantes ou au niveau local.

AU NIVEAU DES PRODUITS				
Fonction	Produits visés	Indicateurs possibles	Sources de vérification	Commentaires, postulats
La gestion globale	L'AAC est indépendante et financièrement solide	Montant des financements reçus par l'AAC	Informations internes	Toute AAC a besoin de financements pour exercer ses activités. Si la collecte de fonds n'entre pas toujours dans la « sphère d'influence » de l'AAC, elle n'en constitue pas moins une tâche indispensable et donc, un indicateur approprié.
La gestion globale	L'AAC propose les formations nécessaires à son personnel	Nombre d'évaluations des besoins de formation entreprises	Rapports d'évaluations des besoins de formation	Lors d'une étape ultérieure, un rapport standardisé d'évaluation des besoins de formation peut mesurer l'amélioration des compétences du personnel.
		Nombre de formations effectuées et finalisées par le personnel de l'AAC	Informations internes	Cet indicateur est assez sommaire dans la mesure où il n'apprécie pas la qualité des formations. Il devrait donc être combiné avec les indicateurs qui suivent.
		Nombre / pourcentage du personnel de l'AAC qui considère avoir acquis de nouvelles compétences	Auto-évaluation guidée effectuée lors des entretiens annuels	Il est possible d'utiliser des formulaires d'auto-évaluation, mais les auto-évaluations guidées permettent d'obtenir des résultats plus fiables.
		Nombre / pourcentage du personnel ayant obtenu un certificat spécialisé / un diplôme ou une formation de niveau master dans le domaine de la lutte contre la corruption	Informations internes	Cet indicateur pourrait être affiné en étant combiné avec les plans de développement du personnel de l'agence.
		Nombre des employés de l'AAC ayant effectué des visites d'étude pour passer en revue les meilleures pratiques de lutte contre la corruption	Informations internes	S'il existe déjà des plans spécifiques de visites d'étude, cet indicateur peut être transformé en indicateur qualitatif, les visites d'étude spécifiques faisant office d'étapes.

AU NIVEAU DES PRODUITS				
Fonction	Produits visés	Indicateurs possibles	Sources de vérification	Commentaires, postulats
La gestion globale	L'AAC dispose de l'effectif adéquat (le nombre approprié d'employés doit être défini en concertation entre l'AAC, les donateurs et le gouvernement, sur la base des fonctions, des tâches et des attentes de l'AAC)	Nombre d'employés	Informations internes	Si l'AAC souffre vraiment d'un manque de personnel, il peut être utile d'évaluer son effectif.
		L'agence compte l'effectif nécessaire (oui/non)	Informations internes	Ce critère peut être utilisé comme indicateur qualitatif si un objectif a été défini en matière de ressources humaines. Il peut être ventilé par étapes à accomplir.
		Taux de recrutement	Informations internes	Le taux de recrutement constitue un autre indicateur sur l'effectif de l'agence.
		Taux de rétention des employés	Informations internes	Au fil du temps, le taux de rétention peut s'avérer plus important que le taux de recrutement.
La gestion globale	Le matériel et les logiciels informatiques sont adaptés aux besoins de l'AAC	Les besoins de l'AAC en termes de matériel et de logiciels informatiques ont été analysés et un plan a été mis en œuvre (oui/non)	Informations internes	Cet indicateur qualitatif peut être ventilé par étapes.
La gestion globale	L'AAC est capable de communiquer sur ses réussites et d'avoir des échanges avec la société	Nombre de visites / téléchargements / durée des visites / échanges réalisés sur le site Internet de l'agence	Informations internes	Cet indicateur est révélateur de l'utilité du site Internet de l'agence dans les pays où l'accès à Internet est généralisé.
		Nombre de réunions d'information et d'événements organisés par l'AAC	Informations internes	Cela peut inclure les rencontres avec des journalistes, des communautés, des organisations de jeunes, etc.

AU NIVEAU DES PRODUITS				
Fonction	Produits visés	Indicateurs possibles	Sources de vérification	Commentaires, postulats
La gestion globale	L'AAC sert l'intérêt public et est réactive	Pourcentage de réponses aux signalements de corruption dans un délai prédéterminé en heures	Informations internes	Peut être utilisé comme indicateur interne et comme facteur incitatif de service aux citoyens.
		Pourcentage de réponses aux demandes de conseils en matière de prévention de la corruption dans un délai prédéterminé en jours ouvrables	Informations internes	Peut être utilisé comme indicateur interne et comme facteur incitatif de service aux citoyens.
La législation	L'AAC contribue à la législation anti-corruption	Nombre de contributions apportées par l'AAC aux projets de lois	Informations internes	On ne peut attendre de l'AAC qu'elle rédige des projets de lois pour le parlement, mais l'indicateur de performance vise à déterminer si elle a effectué des contributions constructives en la matière.
	L'AAC entreprend des vérifications de conformité de la législation anti-corruption	Nombre de vérifications de conformité effectuées	Informations internes	Cet indicateur pourrait être ventilé par secteurs ou par conventions internationales.
La répression des infractions	L'AAC dispose d'un système facilitant le traitement des signalements de corruption	Nombre de signalements de corruption reçus	Informations internes	Cet indicateur peut être ventilé (comme dans l'exemple de l'ICAC que nous avons vu plus haut). Il peut être utile de distinguer les signalements relatifs à la petite corruption de ceux qui concernent la grande corruption.
	L'AAC est en mesure d'enquêter efficacement sur des affaires de corruption	Nombre d'enquêtes en cours	Informations internes	Indicateur d'accomplissement en matière d'enquêtes et non pas de résultats.
La prévention	L'AAC promeut un sens aigu du service public	Nombre de fonctionnaires formés sur les questions de conflits d'intérêts, d'éthique et sur les mesures anti-corruption internes	Informations internes	Mesure purement quantitative du nombre de fonctionnaires formés. Si plusieurs types de formations sont proposées, l'indicateur peut être ventilé selon leur genre.

AU NIVEAU DES PRODUITS				
Fonction	Produits visés	Indicateurs possibles	Sources de vérification	Commentaires, postulats
La prévention	L'AAC contribue à l'élaboration et à la mise en œuvre de codes de conduite	Nombre d'institutions publiques ayant élaboré et mis en œuvre des codes de conduite	Informations internes	On part du principe que l'AAC remplit une fonction de conseils en matière de codes de conduite.
La prévention	L'AAC contribue à l'élaboration et à la mise en œuvre de régimes d'intégrité publique	Nombre d'institutions publiques ayant élaboré et mis en œuvre des régimes d'intégrité	Informations internes	On part du principe que l'AAC remplit une fonction de conseils en matière de régimes d'intégrité.
La prévention	Le patrimoine des fonctionnaires fait l'objet de déclarations et d'analyses	Nombre de déclarations de patrimoine analysées	Informations internes	Indicateur sommaire, et un nombre élevé de déclarations analysées ne constitue pas nécessairement un indice d'efficacité. Les indicateurs les plus importants en matière de patrimoine sont les indicateurs de résultats (voir plus haut).
La prévention	L'AAC contribue à l'amélioration des audits	Pourcentage du budget du gouvernement audité selon les normes requises au cours de l'année budgétaire	Statistiques du bureau du vérificateur des comptes	L'AAC n'est généralement pas directement responsable des audits, mais si c'est une question prioritaire et que l'agence œuvre à l'améliorer en collaboration avec d'autres institutions, l'indicateur est alors valide. On ne peut toutefois attribuer aucun effet de causalité à l'AAC.
		Pourcentage / nombre d'audits ayant révélé des irrégularités	Statistiques du bureau du vérificateur des comptes	On part du principe que l'AAC contribue à améliorer les audits par le biais du renforcement des capacités des institutions.
	L'AAC contribue à l'amélioration des procédures d'achats publics	Pourcentage de contrats et de marchés publics contrôlés au cours de l'année budgétaire	Statistiques du gouvernement	Les achats publics relèvent rarement de la responsabilité des AAC, mais si de réels efforts sont consacrés à l'amélioration du travail dans ce domaine, cet indicateur est alors utile. On ne peut toutefois attribuer aucun effet de causalité à l'AAC.

AU NIVEAU DES PRODUITS				
Fonction	Produits visés	Indicateurs possibles	Sources de vérification	Commentaires, postulats
		Pourcentage / nombre d'inspections de contrats et de marchés publics effectuées par le gouvernement et ayant révélé des irrégularités	Statistiques du gouvernement	Comme précédemment, seul cet indicateur cible le résultat des contrôles effectués.
La prévention	L'AAC sensibilise le public aux conséquences négatives de la corruption	Campagne nationale de sensibilisation planifiée et lancée (oui/non)	Informations internes	Indicateur qualitatif qui peut être ventilé par étapes.
		Nombre d'ateliers et de séminaires organisés pour les journalistes, les ONG et l'administration publique dans le cadre du mandat de l'AAC	Informations internes	Cet indicateur peut être ventilé, par exemple, par type de public cible.
La coopération entre agences	L'AAC assure la promotion de la coopération entre les institutions impliquées dans la lutte contre la corruption	Nombre de conférences réalisées	Informations internes	Ce critère peut devenir un indicateur qualitatif en indiquant quand et comment les conférences ont lieu.
	L'AAC organise des ateliers intergouvernementaux communs	Nombre d'ateliers intergouvernementaux communs réalisés	Informations internes Comptes-rendus des rencontres	L'organisation régulière d'ateliers communs entre les institutions impliquées dans la lutte contre la corruption contribue à améliorer la coordination et la coopération.
La production et la gestion des connaissances	L'AAC a les capacités de gérer un système de S&E	Présence d'un système de S&E (oui/non)	Informations internes Evaluation d'expert externe	Indicateur qualitatif qui peut être ventilé en sous-indicateurs.
	L'AAC collecte ses données de manière systématique et régulière	Présence d'une personne chargée de la collecte et du stockage des informations de suivi (oui/non)	Informations internes	Cette personne devrait normalement être un membre du personnel du service de S&E.

AU NIVEAU DES PRODUITS				
Fonction	Produits visés	Indicateurs possibles	Sources de vérification	Commentaires, postulats
		Les données sont régulièrement compilées pour tous les indicateurs choisis (oui/non)	Informations internes	L'AAC doit définir la « régularité » en fonction du contexte et de ses ressources humaines. Elle peut par exemple définir une compilation trimestrielle ou semestrielle des données.
	L'AAC effectue des recherches et des études analytiques sur la corruption, telles que des évaluations de risques	Nombre de rapports de recherches produits	Informations internes	Cet indicateur ne mesure pas la qualité des rapports de recherches. On part du principe que les rapports de mauvaise qualité ne sont pas validés.
		Nombre de rapports d'évaluation de risques de corruption produits	Informations internes	Cet indicateur se concentre sur l'évaluation des risques, mais pourrait cibler d'autres champs du travail analytique.
La coopération internationale	L'AAC contribue aux échanges internationaux d'expertise	Nombre de demandes internationales adressées à l'AAC pour des services d'expertise à l'étranger	Informations internes	Le nombre de demandes adressées à l'AAC par des institutions similaires souhaitant tirer un apprentissage des pratiques et du personnel de l'agence traduit ici la réputation professionnelle de cette dernière.
		Nombre de délégations d'AAC étrangères reçues	Informations internes	Si des délégations d'institutions similaires prennent la peine de se déplacer jusqu'à l'AAC dans une optique d'apprentissage, on peut supposer que l'agence jouit d'une excellente réputation professionnelle.
		Nombre de visiteurs étrangers à l'AAC	Informations internes	Cet indicateur mesure les visites individuelles plutôt qu'institutionnelles, et inclut les professionnels, les chercheurs, le personnel des organismes donateurs, etc.
La coopération avec la société civile	L'AAC fournit des informations / des activités pédagogiques à la société civile	Nombre de rencontres / ateliers organisés avec des journalistes, des leaders de groupes de jeunes, des représentants d'ONG, etc.	Informations internes	Le terme de rencontre reste large. Cet indicateur mesure le travail de proximité effectué par l'AAC.

AU NIVEAU DES PRODUITS

Fonction	Produits visés	Indicateurs possibles	Sources de vérification	Commentaires, postulats
	Les organisations de la société civile (ONG, écoles, syndicats, etc.) sollicitent et adoptent les supports de l'AAC	Nombre d'organisations de la société civile ayant sollicité et adopté des codes de conduite, des guides éthiques, etc.	Informations internes	Cet indicateur dépend des initiatives et des supports proposés par l'AAC.
La coopération avec le secteur privé	L'AAC fournit des informations / des activités pédagogiques aux entreprises sur la prévention de la corruption	Nombre de rencontres / ateliers organisés avec des entreprises, des associations professionnelles, etc.	Informations internes	Le terme de rencontre reste large. Cet indicateur mesure le travail de proximité effectué par l'AAC.
	Les entreprises sollicitent et adoptent les codes de conduite, les guides éthiques et les autres supports recommandés par l'AAC	Nombre d'entreprises ayant sollicité et adopté des codes de conduite, des guides éthiques, etc.	Informations internes	Cet indicateur dépend des initiatives et des supports proposés par l'AAC.

Bibliographie

- Arndt (Christine) et Oman (Charles), *Uses and Abuses of Governance Indicators*, Paris, Centre de développement de l'OCDE, 2006.
- Bamberger (Michael), *Conducting Quality Impact Evaluations under Budget, Time and Data Constraints*, Washington, DC, Groupe indépendant d'évaluation de la Banque mondiale, 2006.
- Bamberger (Michael), Rao (Vijaendra) et Woolcock (Michael), "Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development", in *Handbook of Mixed Methods Research*, Thousand Oaks, CA, Sage, édité par Abbas Tashakkori et Charles B. Teddlie, 2ème éd., 2010, 613–42.
- Bamberger (Michael), Rugh (Jim) et Mabry (Linda), *Real World Evaluation: Working under Budget, Time, Data and Political Constraints*, Thousand Oaks, CA, Sage, 2006.
- Bamberger (Michael) et White (Howard), "Using Strong Evaluation Designs in Developing Countries: Experiences and Challenges", *Journal of Multidisciplinary Evaluation* 4, n° 8, octobre 2007, 58–73.
- Banerjee (Abhijit), Rema (Hanna) et Mullainathan (Sendhil), "Corruption", http://www.hks.harvard.edu/fs/rhanna/Corruption_Chapter_Full8.pdf, 2009.
- Bardhan (Pranab), "Corruption and Development: A Review of Issues", *Journal of Economic Literature* 35, septembre 1997, 1320–46.
- Bjorkman (Martina) et Svensson (Jacob), "Power to the People: Evidence from a Randomized Field Experiment of a Community-Based Monitoring Project in Uganda", *Quarterly Journal of Economics* 124, n° 2, 2009, 735–69.
- Boardman (Anthony E.), Greenberg (David H.), Vining (Aidan R.) et Weimer (David L.), *Cost-Benefit Analysis*, Upper Saddle River, NJ, Pearson International, 3ème éd., 2006.
- Bryman (Alan), *Social Research Methods*, Oxford, Royaume-Uni, Oxford University Press, 2ème éd, 2004.
- Byrne (Elaine), Arnold (Anne-Katrin) et Nagano (Fumiko), *Building Public Support for Anti-Corruption Efforts: Why Anti-Corruption Agencies Need to Communicate and How*, Washington, DC, Banque mondiale, 2010.
- Center for Global Development, *When Will We Ever Learn? Improving Lives through Impact Evaluation*, Washington, DC, Center for Global Development, 2006.
- Commission européenne, *Project Cycle Management Guidelines*, Bruxelles, Office de coopération de la Commission européenne EuropeAid, 2004.
- , *Cause-and-Effect Analysis*, Bruxelles, Commission européenne, http://ec.europa.eu/europeaid/evaluation/methodology/methods/mth_att_en.htm, 2006a.
- , *Evaluation Methods for the European Union's External Assistance*, vol. 1, *Methodological Bases for Evaluation*, Bruxelles, Commission européenne, 2006b.
- , *Evalsed: The Resource for the Evaluation of Socio-Economic Development*, Sourcebook 2, *Methods & Techniques—Evaluability Assessment*, Bruxelles, Commission européenne, http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/sourcebooks/method_techniques/structuring_evaluations/evaluability/index_en.htm, 2009.
- Davies (Rick) et Dart (Jess), *The 'Most Significant Change' (MSC) Technique: A Guide to Its Use*, Cambridge, Royaume-Uni, Monitoring and Evaluation News, <http://www.mande.co.uk/docs/MSCGuide.htm>, 2005.
- Denzin (Norman K.), *The Research Act in Sociology*, London, Butterworths, 1970.

- De Sousa (Luís), *Anti-Corruption Agencies: Between Empowerment and Irrelevance*, Florence, Institut universitaire européen, 2009a.
- , *Does Performance Matter to Institutional Survival? The Method and Politics of Performance Measurement for Anti-Corruption Agencies*, Florence, Institut universitaire européen, 2009b.
- De Sousa (Luís), Hindness (Barry) et Larmour (Peter), *Governments, NGOs and Anti-Corruption: The New Integrity Warriors*. London, Routledge, éd. 2009.
- De Speville (Bertrand), “Why Do Anti-Corruption Agencies Fail?”, in “Implementation Tools: Inputs for a United Nations Experts Group Meeting”, Vienne, (non publié), Programme mondial de lutte contre la corruption (GPAC) et Centre des Nations unies pour la prévention de la criminalité internationale (UNCICP), 2000.
- DFID (Ministère britannique du Développement international), *Guidance on Using the Revised Logical Framework*, How to Note: DFID Practice Paper, Londres, DFID, <http://www.dfid.gov.uk/Documents/publications1/how-to-guid-rev-log-fmwk.pdf>, 2009.
- Doig (Alan), “Good Government and Sustainable Anti-Corruption Strategies: A Role for Independent Anti-Corruption Agencies”, *Public Administration and Development* 15, n° 2, 1995, 151–65.
- , “Getting the Boring Bits Right First: Capacity Building for Anti-Corruption Agencies (ACA)”, in “Implementation Tools: Inputs for a United Nations Experts Group Meeting”, Vienne, (non publié), Programme mondial de lutte contre la corruption (GPAC) et Centre des Nations unies pour la prévention de la criminalité internationale (UNCICP), 2000.
- , “Markets, Management, Managed Work and Measurement: Setting the Context for Effective Anti-Corruption Commissions”, in *Governments, NGOs and Anti-Corruption: The New Integrity Warriors*, Londres, édité par Luís De Sousa, Barry Hindness et Peter Larmour, Routledge, 2009.
- Doig (Alan), Watt (David) et Robert Williams, *Measuring ‘Success’ in Five African Anti-Corruption Commissions: The Cases of Ghana, Malawi, Tanzania, Uganda & Zambia*, Rapport U4, Bergen, Norvège, U4 Anti-Corruption Centre, 2005.
- , “Why Developing Country Anti-Corruption Agencies Fail to Develop: Understanding the Dilemmas of Organisational Development, Performance Expectation and Donor and Government Cycles in African Anti-Corruption Agencies”, *Public Administration and Development* 27, n° 3, 2007, 251–59.
- Duncan (Nick), “The Non-Perception Based Measurement of Corruption: A Review of Issues and Methods from a Policy Perspective.” In *Measuring Corruption*, Farnham Surrey, Royaume-Uni, Ashgate, édité par Charles Sampford, Arthur Shacklock, Carmel Connors et Fredrik Galtung, 2006, 131–62.
- Earl (Sarah), Carden (Fred) et Smutylo (Terry), *Outcome Mapping: Building Learning and Reflection into Development Programs*, Ottawa, Centre de recherches pour le développement international (CRDI), 2001.
- Foresti (Marta), Sharma (Bhavna), O’Neil (Tammie) et Evans (Alison), *Evaluation of Citizens’ Voice and Accountability: Evaluation Framework*, Londres, Overseas Development Institute, 2007.
- Galtung (Fredrik), “Measuring the Immeasurable: Boundaries and Functions of (Macro) Corruption indices”, in *Measuring Corruption*, Farnham Surrey, Royaume-Uni, Ashgate, édité par Charles Sampford, Arthur Shacklock, Carmel Connors et Fredrik Galtung, 2005, 101–30.
- Gertler (Paul J.), Martinez (Sebastian), Premand (Patrick), Rawlings (Laura B.) et Vermeersch (Christel M. J.), *Impact Evaluation in Practice*, Washington, DC, Banque mondiale, 2011.
- Gittinger (J. Price), *Economic Analysis of Agricultural Projects*, Baltimore, Johns Hopkins University Press, 2ème éd., 1982.
- Glasson (John), Therivel (Riki) et Chadwick (Andrew), *Introduction to Environmental Impact Assessment*, Londres, Routledge, 3ème éd., 2005.

- Groupe de travail des Partenaires européens contre la corruption (EPAC), *Common Standards and Best Practices for Anti-Corruption Agencies*, http://www.stt.lt/documents/tarptautinis_bendradarbiavimas/KNAB_elektroniskais_buklets.pdf, 2008.
- ICAC (Commission indépendante contre la corruption, Région administrative spéciale de Hong Kong), *2009 Annual Report*, Hong Kong, ICAC, http://www.icac.org.hk/filemanager/en/Content_1020/2009.pdf, 2009.
- , *Performance Pledge*, Hong Kong, ICAC, http://www.icac.org.hk/filemanager/en/Content_1017/pledge_eng.pdf, 2011.
- IEG (Groupe indépendant d'évaluation), *Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank*, Washington, DC, Banque mondiale, 2006.
- KNAB (Bureau de prévention et de lutte contre la corruption en Lettonie), *Progress and Results in Preventing and Combating Corruption in Latvia: Periodical Update*, Riga, KNAB, http://www.knab.gov.lv/uploads/eng/periodic_update_20091.pdf, 2009.
- Knack (Stephen), *Measuring Corruption in Eastern Europe and Central Asia: A Critique of the Cross-Country Indicators*, Policy Research Working Paper 3968, Washington, DC, Banque mondiale, 2006.
- Langseth (Petter), "Measuring Corruption", in *Measuring Corruption*, édité par Charles Sampford, Arthur Shacklock, Carmel Connors et Fredrik Galtung, Farnham Surrey, Royaume-Uni, Ashgate, 2006, 7–44.
- Lusthaus (Charles), Adrien (Marie-Hélène), Anderson (Gary), Carden (Fred) et Plinio Montalván (George), *Organizational Assessment: A Framework for Improving Performance*, Ottawa, Centre de recherches pour le développement international, Washington, DC, Banque interaméricaine de développement, 2002.
- Lusthaus (Charles), Anderson (Gary) et Murphy (Elaine), *Institutional Assessment: A Framework for Strengthening Organisational Capacity of IDRC's Research Partners*, Ottawa, Centre de recherches pour le développement international, 1995.
- McGee (Rosemary) et Gaventa (John), "Review of Impact and Effectiveness of Transparency and Accountability Initiatives", Document élaboré pour un Atelier de Transparency and Accountability Initiative, octobre 2010, 14–15.
- Meagher (Patrick) *Anti-Corruption Agencies: A Review of Experience*, Washington, DC, Banque mondiale, 2002.
- , *Anti-Corruption Agencies: A Review of Experience*. IRIS Discussion Paper 04/02. College Park, MD, University of Maryland Center for Institutional Reform and the Informal Sector, 2004.
- Morra (Imas), Rist (Linda G. et Ray C.), *The Road to Results: Designing and Conducting Effective Development Evaluations*, Washington, DC, Banque mondiale, 2009.
- OCDE (Organisation de coopération et de développement économiques), *Ethics in the Public Service: Current Issues and Practice*, Hors-série de PUMA, Paris, OCDE, 1996.
- OCDE/CAD (Organisation de coopération et de développement économiques, Comité d'aide au développement), *Methods and Procedures in Aid Evaluation*, Paris, OCDE, 1986.
- , *Principles for Evaluation of Development Assistance*, Paris, OCDE, 1991.
- , *Glossary of Evaluation and Results-Based Management (RBM) Terms*, Paris, OCDE, 2000.
- , *Evaluation Feedback for Effective Learning and Accountability*, Paris, OCDE, 2001.
- , *Glossary of Terms in Evaluation and Results-Based Management*, Paris, OCDE, 2002.

- , “DAC Evaluation Quality Standards (for test phase application)”, Avant-projet, Réseau CAD sur l’Evaluation du développement, Paris, OCDE, <http://siteresources.worldbank.org/EXTGLOREGPARPROG/Resources/DACEvaluationQualityStandards.pdf>, 2006.
- , *Specialised Anti-Corruption Institutions: Review of Models*, Paris, OCDE, 2007.
- , *Glossary of Key Terms in Evaluation and Results-Based Management*, Paris, OCDE, 2009.
- , *Evaluating Development Co-operation: Summary of Key Norms and Standards*, Paris, Réseau CAD sur la Coopération pour le développement (OCDE), 2ème éd., 2010a.
- , *Quality Standards for Development Evaluation*, Lignes directrices et ouvrages de référence du CAD, Réseau CAD sur l’Evaluation du développement, Paris, OCDE, <http://www.oecd.org/dataoecd/30/62/36596604.pdf>, 2010b.
- , “DAC Criteria for Evaluating Development Assistance”, Direction de la Coopération pour le développement (OCDE), http://www.oecd.org/document/22/0,2340,en_2649_34435_2086550_1_1_1_1,00.html, 2011.
- Olken (Benjamin A.), “Monitoring Corruption: Evidence from a Field Experiment in Indonesia”, *Journal of Political Economy* 115, n° 2, 2007, 200–49.
- Osborne (David) et Gaebler (Ted), *Reinventing Government*. New York, Penguin, 1992.
- Picciotto (Robert), “The Value of Evaluation Standards: A Comparative Assessment”, *Journal of Multidisciplinary Evaluation*, n° , octobre 2005, 30–59.
- PNUD (Programme des Nations unies pour le développement), *Institutional Arrangements to Combat Corruption: A Comparative Study*, Bangkok, PNUD, 2005.
- , *A User’s Guide to Measuring Corruption*, Oslo, Centre d’Oslo sur la Gouvernance (PNUD), 2008.
- , *Practitioners’ Guide: Capacity Assessment of Anti-Corruption Agencies (ACAs)*, New York: PNUD, 2011.
- Pope (Jeremy), *The Need for, and Role of, an Independent Anti-Corruption Agency*, Documents de travail de Transparency International, Berlin, Transparency International, 1999.
- Pope (Jeremy) et Vogl (Frank), “Making Anti-Corruption Agencies More Effective”, *Finance and Development* 37, n° 2, juin 2000, 6–9.
- Quah (Jon), “Accountability and Anticorruption Agencies in the Asia-Pacific Region”, in *Combating Corruption in Asian and Pacific Economies*, Manille, Banque asiatique de développement, 2000, 101–24.
- Ravallion (Martin), “Evaluating Three Stylised Interventions”, *Journal of Development Effectiveness* 1, n° 3, septembre 2009, 227–36
- Ravindra (Adikeshavalu), *An Assessment of the Impact of Bangalore Citizen Report Cards on the Performance of Public Agencies*, Document de travail sur l’ECD 12, Washington, DC, Département de l’évaluation des opérations, Banque mondiale, 2004.
- Recovery & Development Consortium, *DFID Yemen Social Fund for Development: Impact Evaluation*, Kongens Lyngby, Danemark, Recovery & Development Consortium, http://www.sfd-yemen.org/SFD_SITE/admin/PDF/upload_DOC/SFD_Impact_Evaluation_Report_2010.pdf, 2010.
- Reinikka (Ritva) et Svensson (Jacob), *Cost Efficiency in Healthcare*, Washington, DC, Banque mondiale, 2000.

- , *Survey Techniques to Measure and Explain Corruption*, Washington, DC, Banque mondiale, 2003.
- Schiavo-Campo (Salvatore), “Performance in the Public Sector.” *Asian Journal of Political Science* 7, n° 2, 1999, 75–87.
- Smilov (Daniel), “Anti-Corruption Bodies as Discourse-Controlling Instruments: Experiences from South East Europe”, in *Governments, NGOs and Anti-Corruption: The New Integrity Warriors*, Londres, Routledge, édité par Luís De Sousa, Barry Hindess et Peter Larmour, 2009.
- Søreide (Tina), *Estimating Corruption*. Bergen, Norvège, Chr. Michelsen Institute, 2003.
- , *Is It Wrong to Rank? A Critical Assessment of Corruption Indices*. Document de travail CMI 2006:1, Bergen, Norvège, Chr. Michelsen Institute, 2006.
- Stake (Robert E.), *The Art of Case Study Research*. Thousand Oaks, CA, Sage, 1995.
- Sundet (Geir), *Following the Money: Do Public Expenditure Tracking Surveys Matter?* U4 Issue 2008:8, Bergen, Norvège, U4 Anti-Corruption Centre, 2008.
- TI et PNUD (Transparency International et Programme des Nations unies pour le développement), *Mapping of Corruption and Governance Measurement Tools in Sub-Saharan Africa*, Berlin, Transparency International, Oslo, Centre d’Oslo sur la Gouvernance (PNUD), 2007.
- UNIFEM (Fonds de développement des Nations unies pour la femme), *Guidance Note on Carrying Out an Evaluability Assessment*, Série Notes indicatives sur l’évaluation, New York, Service d’Evaluation de l’UNIFEM, 2009.
- University of London, *Project Appraisal and Impact Analysis Course Reader*, Londres, Centre for Financial and Management Studies, School of Oriental and African Studies, University of London, 2010.
- UNODC (Office des Nations unies contre la drogue et le crime), *Manuel pratique de lutte contre la corruption du Programme mondial des Nations unies contre la corruption*, Vienne, UNODC, 2002.
- USAID (Agence nord-américaine pour le développement international), *Anti-Corruption Agencies (ACAs)*. Office of Democracy and Governance, Anti-Corruption Programme Brief, Washington, DC, USAID, 2006.
- Wade (Robert), “The System of Administrative and Political Corruption: Canal Irrigation in South India”, *Journal of Development Studies* 18, n° 3, 1982, 287–328.
- , “The Market for Public Office: Why the Indian State Is Not Better at Development”, *World Development* 13, n° 4, 1985, 467–97.
- Webb (Eugene J.), Campbell (Donald T.), Schwartz (Richard D.) et Sechrest (Lee), *Unobtrusive Measures: Nonreactive Measures in the Social Sciences*, Chicago, Rand McNally, 1966.
- White (Howard), *Theory-Based Impact Evaluation: Principles and Practice*, Document de travail 3 de 3ie, New Delhi, International Initiative for Impact Evaluation, 2009.
- Wholey (Joseph S.), Hatry (Harry P.) et Newcomer (Kathryn E.), éd., *Handbook of Practical Programme Evaluation*, San Francisco, Jossey-Bass, 2004.
- Yin (Robert K.), *Case Study Research: Designs and Methods*, Thousand Oaks, CA, Sage, 2003.

U4 Anti-Corruption Resource Centre
Chr. Michelsen Institute (CMI)
Téléphone: +47 47 93 80 00
Fax: +47 47 93 80 01
u4@u4.no
www.U4.no

P.O.Box 6033 Bedriftssenteret
N-5892 Bergen, Norway
Adresse:
Jekteviksbakken 31, Bergen

Cette publication est disponible sur notre site:
<http://www.u4.no/publications-2-fr-FR/>

MOTS CLES:

Agences anti-corruption
Suivi et évaluation
Normes d'évaluation
Evaluation d'impact
Indicateur
Théorie du changement
Garantie de qualité

Photo de couverture:
Patrick Q
patrick_q@flickr.com

Le nombre d'agences anti-corruption (AAC) à travers le monde est en constante augmentation. Pourtant, donateurs internationaux et gouvernements nationaux remettent de plus en plus en question le rôle de ces organismes, qui ne répondraient pas à l'attente générale.

Pour tenter de sonder la solidité des hypothèses formulées au sujet de l'efficacité des AAC, nous avons entrepris de rassembler et d'analyser les évaluations menées par des agences individuelles. Nous nous sommes malheureusement heurtés à un surprenant manque de documentation, le peu d'évaluations ayant été effectuées ne mesurant par ailleurs ni les résultats ni les impacts des AAC. Ces dernières sont exposées à de nombreuses critiques, lesquelles sont rarement justifiées par des arguments concrets. Les preuves devraient être recherchées du côté des AAC elles-mêmes. Il leur faudrait s'employer à établir des indicateurs basés sur les résultats pour démontrer la portée de leurs activités, et collecter toutes les informations nécessaires.

Pour l'heure, il est donc impossible de déterminer si l'échec prétendu des AAC relève d'un problème d'évaluation ou de conception. La valeur de ces organismes ne pourra en effet être établie que sur la base d'évaluations tangibles.

Dans cette optique, le présent document vise à fournir des informations techniques, méthodologiques et pratiques au personnel des AAC souhaitant entreprendre des activités de suivi et d'évaluation ainsi qu'à démontrer comment les résultats et l'impact du travail de ces agences peuvent être évalués à la lumière de critères factuels et objectifs.