



Une application réfléchie des critères d'évaluation



Une application réfléchie des critères d'évaluation

Cet ouvrage est publié sous la responsabilité du Secrétaire général de l'OCDE. Les opinions et les arguments exprimés ici ne reflètent pas nécessairement les vues officielles des pays Membres de l'OCDE.

Ce document, ainsi que les données et cartes qu'il peut comprendre, sont sans préjudice du statut de tout territoire, de la souveraineté s'exerçant sur ce dernier, du tracé des frontières et limites internationales, et du nom de tout territoire, ville ou région.

Les données statistiques concernant Israël sont fournies par et sous la responsabilité des autorités israéliennes compétentes. L'utilisation de ces données par l'OCDE est sans préjudice du statut des hauteurs du Golan, de Jérusalem-Est et des colonies de peuplement israéliennes en Cisjordanie aux termes du droit international.

Note de la République de Türkiye

Les informations figurant dans ce document qui font référence à « Chypre » concernent la partie méridionale de l'île. Il n'y a pas d'autorité unique représentant à la fois les Chypriotes turcs et grecs sur l'île. La Türkiye reconnaît la République Turque de Chypre Nord (RTCN). Jusqu'à ce qu'une solution durable et équitable soit trouvée dans le cadre des Nations Unies, la Türkiye maintiendra sa position sur la « question chypriote ».

Note de tous les États de l'Union européenne membres de l'OCDE et de l'Union européenne

La République de Chypre est reconnue par tous les membres des Nations Unies sauf la Türkiye. Les informations figurant dans ce document concernent la zone sous le contrôle effectif du gouvernement de la République de Chypre.

Merci de citer cet ouvrage comme suit :

OCDE (2023), *Une application réfléchie des critères d'évaluation*, Éditions OCDE, Paris, <https://doi.org/10.1787/716baf6e-fr>.

ISBN 978-92-64-52037-0 (imprimé)

ISBN 978-92-64-32295-0 (pdf)

ISBN 978-92-64-32231-8 (HTML)

ISBN 978-92-64-72866-0 (epub)

Crédits photo : Couverture © TFK.

Les corrigenda des publications sont disponibles sur : www.oecd.org/fr/apropos/editionsocde/corrigendadepublicationsdelocde.htm.

© OCDE 2023

L'utilisation de ce contenu, qu'il soit numérique ou imprimé, est régie par les conditions d'utilisation suivantes : <https://www.oecd.org/fr/conditionsdutilisation>.

Avant-propos

L'évaluation joue un rôle central pour éclairer la conception et la mise en œuvre de politiques et de programmes de développement qui permettent d'obtenir de meilleurs résultats, plus équitables et plus durables. Les données issues de l'évaluation et la réflexion critique que cette dernière peut favoriser sont essentielles pour aider les décideurs et les différents acteurs à faire en sorte que ces politiques et programmes se traduisent par des résultats positifs et durables pour les populations et la planète.

Dans le but de promouvoir la pratique de l'évaluation, le Comité d'aide au développement (CAD) de l'Organisation de coopération et de développement économiques (OCDE) a publié, en 1991, les définitions de cinq critères d'évaluation. Depuis 30 ans, ces critères orientent la conception et l'évolution des évaluations dans le domaine du développement international. Ils sont largement reconnus pour le rôle qu'ils jouent dans l'amélioration de la qualité des pratiques d'évaluation et leur caractère collaboratif. Ils ont permis à une diversité d'organisations de concevoir et mener des évaluations qui correspondent bien aux besoins des décideurs, et dont les résultats, qu'ils soient escomptés ou bien inattendus, apportent des données et des éclairages utiles.

Les cinq critères du CAD ont fait l'objet d'une révision en 2018-19. Le processus a mis à profit près de trois décennies d'apprentissage par les membres du Réseau du CAD sur l'évaluation du développement (EvalNet) et la communauté mondiale de l'évaluation en général. Il s'appuyait sur le Programme de développement durable à l'horizon 2030, notamment les Objectifs de développement durable (ODD), et sur l'Accord de Paris dans le cadre de la Convention-cadre des Nations Unies sur les changements climatiques (CCNUCC). Une nouvelle série de définitions, assortie de deux principes d'utilisation, a été publiée en décembre 2019.

Les présentes Orientations complètent les définitions des critères et leurs principes d'utilisation. Elles visent à favoriser une application réfléchie des critères, adaptée aux différents contextes. Augmentées de questions et d'exemples, les Orientations encouragent la réflexion critique et une analyse nuancée. Elles aideront les lecteurs dans la recherche de moyens d'interpréter et d'appliquer les critères afin d'améliorer la qualité de leurs évaluations et de promouvoir plus efficacement l'apprentissage et l'exercice de la redevabilité.

L'apprentissage relatif à l'utilisation optimale de ces critères révisés – dans la coopération internationale pour le développement mais aussi dans d'autres domaines – n'en est qu'à ses prémices. Le but est de continuer de recueillir enseignements et éclairages au fil de leur utilisation et de revoir ces Orientations à la lumière de l'expérience collective et du retour d'information de la communauté mondiale de l'évaluation. À cet égard, les leçons tirées des nouvelles façons d'utiliser les critères en tenant compte des principes du Programme 2030 seront particulièrement utiles – par exemple, les évaluations participatives ou centrées sur les droits de la personne, les évaluations de processus de changement complexes, ou encore celles qui prennent en compte les synergies et les arbitrages selon une approche globale.

Remerciements

Le Secrétariat tient à remercier la communauté mondiale de l'évaluation pour la richesse des discussions et des débats consacrés aux critères au cours de la période 2017-19, lesquels ont contribué à établir la nécessité d'élaborer ces Orientations et à en définir le contenu. En particulier, les participants à la conférence sur les Capacités nationales d'évaluation (CNE), le Groupe des Nations Unies pour l'évaluation (GNUE) et le Groupe de coopération pour l'évaluation (ECG) ont été des partenaires clés dans le processus d'adaptation des critères, et nombre de leurs membres ont apporté des contributions qui sont prises en compte dans ce document. Le Secrétariat exprime sa reconnaissance aux partenaires d'évaluation du monde entier pour leurs contributions et se réjouit à la perspective d'autres discussions pour développer davantage ces Orientations, qui constituent pour l'heure un document de travail.

Les Orientations prennent en compte les nombreuses contributions des membres du Réseau sur l'évaluation du développement (EvalNet), un organe subsidiaire du Comité d'aide au développement (CAD) de l'Organisation de coopération et de développement économiques (OCDE). Les membres d'EvalNet ont apporté maints commentaires et éclairages utiles au sujet des projets de définitions des critères, et ont eu des débats approfondis sur les notions présentées dans ce document – processus qui, souvent, a nécessité de longues discussions au sein des organismes et ministères membres, ainsi que des consultations avec les parties prenantes de l'évaluation du développement. Les membres d'EvalNet ont partagé leurs propres expériences et difficultés, exemples concrets à l'appui, ce qui a permis de donner corps au document. Le Groupe de travail d'EvalNet sur l'égalité femmes-hommes a contribué à l'élaboration des éléments portant sur cette thématique. Le Secrétariat remercie également vivement le Président d'EvalNet, Per Øyvind Bastøe, pour le rôle central qu'il a joué en guidant le processus d'adaptation et l'élaboration de ces Orientations.

Pour l'OCDE, Megan Grace Kennedy-Chouane a piloté le processus d'élaboration de ces Orientations, avec le concours et la contribution de Rahul Malhotra, Hans Lundgren, Alison Pollard, Ola Kasneci et Yee King Ho. De nombreux autres collègues de la Direction de la coopération pour le développement ont formulé des observations sur les versions précédentes du document. Andrew Brenner et Suzanne Parandian ont apporté des commentaires d'ordre éditorial. Stephanie Coic et TFK se sont chargés de la conception graphique des images et des graphiques.

Ima Bishop, Joe Watkins et Nick York (IOD PARC) ont rédigé le projet d'Orientations et apporté de précieuses contributions à son contenu. Julia Betts a revu le document et formulé des observations en retour, après avoir largement contribué au processus de révision des critères initiaux et à l'élaboration et la conceptualisation des nouvelles définitions.

La version française a été traduite par l'OCDE. Le Secrétariat remercie également l'Évaluation indépendante du développement (IDEV) de la Banque africaine de développement et Mirianaud Oswald Agbadome pour la révision et la finalisation du document.

Table des matières

| | |
|---|-----------|
| Avant-propos | 3 |
| Remerciements | 4 |
| Sigles et abréviations | 8 |
| Résumé | 10 |
| 1 Objet et utilité des Orientations : de meilleurs critères pour de meilleures évaluations | 12 |
| Raison d'être des Orientations | 13 |
| Objet des Orientations et mode d'emploi | 13 |
| Utilisateurs cibles des Orientations | 15 |
| Références | 16 |
| Notes | 16 |
| 2 Les six critères : finalité et rôle dans l'évaluation | 17 |
| Qu'est ce qu'un critère d'évaluation ? | 18 |
| Quels types d'interventions peuvent être évalués à l'aune de ces critères ? | 19 |
| Quelle est la place des critères dans le domaine de l'évaluation en général ? | 20 |
| Références | 21 |
| Notes | 21 |
| 3 Utiliser les critères d'évaluation dans la pratique | 22 |
| Une application réfléchie des critères | 23 |
| Adapter les critères à la finalité de l'évaluation | 24 |
| Comprendre la dimension temporelle | 26 |
| Utiliser les critères pour améliorer les questions d'évaluation | 27 |
| Liens entre les critères | 28 |
| Sélectionner les critères à utiliser | 28 |
| Mettre en œuvre et adapter les critères au niveau institutionnel | 30 |
| Répondre aux exigences du Programme de développement durable à l'horizon 2030 et des Objectifs de développement durable | 30 |
| Appliquer un prisme sexospécifique aux critères | 33 |
| Utilisation d'autres critères | 34 |
| Références | 35 |
| Notes | 36 |

| | |
|--|-----------|
| 4 Comprendre les six critères : définitions, éléments d'analyse et principaux défis | 38 |
| Pertinence | 40 |
| Cohérence | 48 |
| Efficacité | 55 |
| Efficience | 61 |
| Impact | 68 |
| Viabilité/durabilité | 76 |
| Références | 83 |
| Notes | 88 |

GRAPHIQUES

| | |
|--|----|
| Graphique 2.1. Comment se placent les critères par rapport aux autres normes et critères ? | 21 |
| Graphique 3.1. Réfléchir à la manière de prendre en compte les critères pour une évaluation donnée | 29 |
| Graphique 4.1. Les six critères et leurs questions associées | 39 |

TABLEAUX

| | |
|---|----|
| Tableau 3.1. Appliquer un prisme sexospécifique aux critères | 34 |
| Tableau 4.1. Difficultés posées par l'évaluation de la pertinence | 45 |
| Tableau 4.2. Difficultés posées par l'évaluation de la cohérence | 51 |
| Tableau 4.3. Difficultés posées par l'évaluation de l'efficacité | 58 |
| Tableau 4.4. Difficultés posées par l'évaluation de l'efficience | 64 |
| Tableau 4.5. Difficultés posées par l'évaluation de l'impact | 72 |
| Tableau 4.6. Difficultés posées par l'évaluation de la viabilité/durabilité | 80 |

ENCADRÉS

| | |
|---|----|
| Encadré 1.1. Autres travaux et publications à venir de l'OCDE sur l'évaluation | 14 |
| Encadré 2.1. Adapter les critères à l'intervention : évaluation du plan d'action du BMZ pour l'inclusion des personnes handicapées | 19 |
| Encadré 3.1. Utilisation réfléchie des critères : évaluation de la dotation de la Norvège en faveur de la société civile dans les pays en développement | 24 |
| Encadré 3.2. Définir la finalité de l'évaluation : exemples de l'Ouganda, du Kenya et de la Suède | 25 |
| Encadré 3.3. S'appuyer sur les ODD pour définir des programmes d'évaluation nationaux | 31 |
| Encadré 3.4. Relier les critères au Programme 2030 et aux ODD | 32 |
| Encadré 4.1. Évaluation de la pertinence des programmes de soutien budgétaire général | 46 |
| Encadré 4.2. Évaluation du projet d'amélioration des moyens de subsistance des agriculteurs dans la « zone sèche » du Myanmar | 47 |
| Encadré 4.3. Évaluation de la pertinence des initiatives en faveur de la santé en Bolivie | 47 |
| Encadré 4.4. Évaluation des efforts déployés par la Norvège pour garantir la cohérence des politiques au service du développement | 52 |
| Encadré 4.5. Évaluation de la cohérence de l'aide à la reconstruction après une catastrophe naturelle aux Philippines | 53 |
| Encadré 4.6. Évaluer la cohérence du programme de coopération pour le développement de la Slovénie avec le Monténégro | 54 |
| Encadré 4.7. Évaluation de l'efficacité et de l'inclusivité du programme australien d'assistance électorale | 59 |
| Encadré 4.8. Évaluer l'efficacité de la stratégie et du programme-pays de la Banque africaine de développement (BAfD) au Cap Vert | 60 |
| Encadré 4.9. Évaluation de l'efficience du soutien de la Banque mondiale en faveur de l'électrification rurale | 65 |
| Encadré 4.10. Évaluer l'efficience des subventions aux intrants agricoles en Afrique subsaharienne | 66 |
| Encadré 4.11. Évaluer l'efficience au niveau des politiques : la politique néerlandaise d'aide au développement visant à améliorer la gestion de l'eau | 67 |
| Encadré 4.12. Évaluer l'efficacité et l'impact du programme « Plus de familles en action » en Colombie | 73 |

| | |
|---|----|
| Encadré 4.13. Évaluer l'impact du soutien apporté par l'Asdi à la consolidation de la paix dans des pays touchés par un conflit ou sortant d'un conflit | 73 |
| Encadré 4.14. Évaluation de l'impact dans le cadre d'une évaluation stratégique de l'action publique | 74 |
| Encadré 4.15. Évaluation de l'impact, de la diffusion et de la montée en puissance d'une approche globale de l'aménagement du territoire aux Philippines | 74 |
| Encadré 4.16. Évaluation de l'impact du programme « Girl Empower » au Libéria | 75 |
| Encadré 4.17. Examen de la viabilité/durabilité dans une évaluation <i>ex post</i> d'un projet japonais de développement rural intercommunal en Afghanistan | 80 |
| Encadré 4.18. Leçons sur la viabilité/durabilité tirées de l'évaluation du retrait du programme de soutien budgétaire général au Malawi, en Ouganda, au Rwanda et en Zambie | 82 |
| Encadré 4.19. Évaluer l'impact et la viabilité/durabilité du programme du DFID en faveur de la santé maternelle | 82 |

Sigles et abréviations

| | |
|--------|--|
| 3ie | International Initiative for Impact Evaluation |
| BaFD | Banque africaine de développement |
| ALNAP | Active Learning Network for Accountability and Performance in Humanitarian Action |
| BMZ | Ministère fédéral allemand de la Coopération économique et du développement |
| CEDAW | Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes |
| CDPH | Convention relative aux droits des personnes handicapées |
| CAD | Comité d'aide au développement |
| Danida | Agence danoise de développement international |
| DEval | Institut allemand pour l'évaluation du développement |
| IFD | Institution de financement du développement |
| DFID | Ministère britannique du Développement international |
| ECG | Groupe de coopération pour l'évaluation |
| UE | Union européenne |
| FAO | Organisation des Nations Unies pour l'alimentation et l'agriculture |
| GIZ | Deutsche Gesellschaft für Internationale Zusammenarbeit |
| ICAI | Commission indépendante sur l'impact de l'aide (Royaume-Uni) |
| IEG | Groupe d'évaluation indépendante, Banque mondiale |
| JICA | Agence japonaise de coopération internationale |
| ONG | Organisation non gouvernementale |
| Norad | Agence norvégienne de coopération pour le développement |
| APD | Aide publique au développement |
| OCDE | Organisation de coopération et de développement économiques |
| ODD | Objectif de développement durable |
| Asdi | Agence suédoise de coopération pour le développement international |
| SMART | Spécifique, mesurable, atteignable, réaliste et défini dans le temps |
| ONU | Organisation des Nations Unies |
| GNUE | Groupe des Nations Unies pour l'évaluation |

| | |
|--------|--|
| CCNUCC | Convention-cadre des Nations Unies sur les changements climatiques |
| HCR | Haut-commissariat des Nations Unies pour les réfugiés |
| PAM | Programme alimentaire mondial |

Résumé

L'Organisation de coopération et de développement économiques (OCDE) a élaboré des définitions communes pour six critères d'évaluation – pertinence, cohérence, efficacité, efficience, impact et viabilité/durabilité – afin de faciliter la conduite d'évaluations cohérentes et de qualité. Ces critères posent un cadre normatif pour déterminer l'intérêt ou la valeur d'une intervention (politique, stratégie, programme, projet ou activité). Ils servent de base de jugement pour l'évaluation.

Les critères peuvent être appréhendés comme un ensemble de prismes apportant des perspectives complémentaires qui, une fois réunies, donnent une vue d'ensemble d'une intervention et de ses résultats. Ils encouragent les évaluateurs – et ceux qui contribuent à la conception ou à la gestion des interventions – à approfondir leur réflexion sur la nature d'une intervention, son processus de mise en œuvre et ses résultats.

Outre les six définitions des critères, deux principes directeurs ont été définis par le Réseau du CAD de l'OCDE sur l'évaluation du développement, à savoir :

1. Principe n°1 : Les critères devraient être appliqués de façon réfléchie pour renforcer la qualité et l'utilité des évaluations.
2. Principe n° 2 : L'utilisation des critères dépend de la finalité de l'évaluation.

Les critères doivent être appliqués de façon réfléchie et adaptés au contexte de l'intervention et de l'évaluation. Même s'ils étaient initialement destinés à être utilisés dans le cadre de la coopération internationale pour le développement, ils peuvent être appliqués dans n'importe quel secteur et pour l'évaluation d'interventions publiques comme privées. Ils peuvent servir à évaluer une grande diversité de domaines et types d'interventions – y compris des stratégies, des politiques et des projets. Différentes méthodologies peuvent être utilisées pour effectuer une évaluation sur la base de ces critères, étant entendu qu'ils ne forment qu'une partie seulement de l'éventail plus large des normes et critères d'évaluation.

Les évaluateurs doivent adopter des méthodes de travail qui prennent en compte de manière réfléchie les différences d'expérience ainsi que la façon dont les dynamiques de pouvoir fondées sur le genre et d'autres formes de discrimination (fondées, par exemple, sur l'âge, l'origine ethnique, le statut social, le handicap) influent sur la mise en œuvre et les résultats de l'intervention.

Les notions fondamentales qui sous-tendent chaque critère sont définies comme suit :

- La pertinence examine dans quelle mesure les objectifs et la conception de l'intervention correspondent aux besoins et aux priorités des bénéficiaires, ainsi que l'alignement sur les politiques et priorités nationales et mondiales, et sur celles des partenaires/institutions. La compréhension de la nature sexuée des dynamiques de pouvoir et la réflexion sur l'engagement de « ne laisser personne de côté » sont essentielles pour appréhender la pertinence d'une intervention. Dans le cas de contextes évolutifs, les évaluations doivent également viser à déterminer si les interventions demeurent pertinentes.

- La cohérence – un critère nouvellement ajouté – examine dans quelle mesure d'autres interventions (en particulier des politiques publiques) appuient ou affaiblissent l'intervention évaluée, et inversement. Il s'agit tant de cohérence interne (au sein d'une institution ou d'une administration), notamment la compatibilité avec les normes et les critères internationaux, que de cohérence externe (avec les interventions menées par d'autres acteurs dans le même contexte). La cohérence recouvre les notions de complémentarité, d'harmonisation et de coordination, ainsi que la mesure dans laquelle l'intervention apporte une valeur ajoutée tout en évitant le chevauchement d'activités. Conformément au Programme de développement durable à l'horizon 2030, une attention plus soutenue doit être accordée à la cohérence, en mettant davantage l'accent sur les synergies (ou les arbitrages) entre les domaines d'action. Ce nouveau critère encourage l'adoption d'une approche intégrée pour comprendre les interventions complexes et leurs résultats.
- L'efficacité examine dans quelle mesure les objectifs et les résultats de l'intervention ont été atteints, ou sont en train de l'être, compte tenu de l'importance relative des objectifs. Cette nouvelle définition encourage l'analyse des résultats différenciés selon les groupes de population et de la mesure dans laquelle l'intervention contribue à réduire les inégalités ou les accentue. L'efficacité est le critère le plus couramment évalué et sert souvent de mesure globale de la réussite d'une intervention.
- L'efficience aide les évaluateurs à examiner dans quelle mesure l'intervention produit, ou est susceptible de produire, des résultats de façon économique et dans les temps. Le terme « économique » désigne la conversion des intrants (fonds, expertise, ressources naturelles, temps, etc.) en résultats, de la façon la plus *économiquement avantageuse* possible, par rapport aux options envisageables dans le contexte. La nouvelle définition inclut la dimension temporelle. Ce critère est l'occasion de vérifier si les ressources consacrées à l'intervention sont justifiées par ses résultats – un aspect qui revêt une importance majeure d'un point de vue pratique et politique. De nombreuses parties prenantes, y compris les bénéficiaires, se soucient de l'efficience car ce facteur peut contribuer à une meilleure utilisation de ressources limitées pour davantage de résultats.
- L'impact est la mesure dans laquelle l'intervention a produit, ou devrait produire, des effets importants et de vaste portée, positifs ou négatifs, intentionnellement ou non. Il s'agit d'évaluer l'importance que l'intervention a in fine ainsi que ses effets potentiellement transformateurs (modification globale et durable des systèmes ou des normes). Le critère d'impact ne se contente pas d'analyser l'efficacité et invite à s'interroger sur la différence que fait réellement l'intervention. C'est au titre de ce critère que les évaluateurs examinent si l'intervention a induit des transformations qui comptent réellement pour les populations.
- La viabilité/durabilité est la mesure dans laquelle les bénéfices nets de l'intervention perdureront ou sont susceptibles de perdurer. Selon le moment où l'évaluation est effectuée, il peut s'agir d'analyser le flux réel de bénéfices nets ou d'estimer la probabilité que les bénéfices nets perdurent à moyen et long terme. Si la notion sous-jacente de bénéfices durables reste la même, la nouvelle définition englobe plusieurs éléments d'analyse – financiers, économiques, sociaux et environnementaux – et l'attention devrait être accordée à l'interaction entre ces éléments.

Diverses difficultés peuvent être rencontrées dans l'application des différents critères, notamment des problématiques courantes liées à une mauvaise conception des interventions ou à la disponibilité des données. Pour aider les évaluateurs et les responsables d'évaluation à surmonter ces difficultés, le document formule des suggestions concrètes, visant notamment à : apprécier « l'évaluabilité » à un stade précoce du processus, collaborer avec les parties prenantes pour consigner les décisions prises ou recréer les données de référence, et faire preuve de transparence quant aux éventuelles limites, de façon à gérer les attentes. Ces propositions, ainsi que le reste des Orientations, seront mises à jour au fil des enseignements tirés de l'expérience.

1 **Objet et utilité des Orientations : de meilleurs critères pour de meilleures évaluations**

Ce chapitre explique la finalité du document et la manière dont il peut aider les lecteurs à mieux appréhender les définitions adaptées des critères et à les appliquer dans le cadre de leurs activités. Il décrit également le processus d'élaboration des Orientations et notamment le rôle qu'ont joué les acteurs de la communauté mondiale de l'évaluation pour éclairer leur conception et leur contenu. Enfin, il se penche sur la manière d'appliquer ces critères de façon réfléchie afin d'améliorer à la fois la réalisation et la conception des évaluations.

Raison d'être des Orientations

Le Comité d'aide au développement (CAD) de l'OCDE a commencé par présenter cinq critères d'évaluation, en 1991, (OCDE, 1991^[1]), avant de les définir plus précisément, en 2002 (OCDE, 2002^[2]). Ces cinq critères – pertinence, efficacité, efficience, impact et viabilité – sont depuis devenus une référence commune pour les évaluateurs de la coopération pour le développement et d'autres domaines.

Dans un souci de prise en compte du Programme de développement durable à l'horizon 2030, et l'occasion se présentant de mettre à profit la vaste expérience acquise et les enseignements tirés depuis la première définition des critères, un vaste processus de consultation et d'adaptation a été lancé en 2017-19 et le CAD a adopté les définitions révisées des critères en décembre 2019 (OCDE, 2019^[3])¹. Au cours du processus de consultation, les utilisateurs ont relevé l'intérêt de nombreux aspects des critères d'évaluation et fait part de leur souhait de continuer à les utiliser pour évaluer leurs futures interventions. Les informations reçues en retour montrent que les critères d'évaluation sont largement utilisés et compris, ce qui favorise la cohérence et un certain niveau de comparabilité². Ces critères ont établi une terminologie commune, simple et neutre ainsi qu'un cadre normatif pour le secteur de l'évaluation qui est adaptable et axé sur les résultats et couvre de manière exhaustive les principaux domaines requis pour l'exercice de la redevabilité et l'apprentissage. De ce fait, les critères révisés traduisent un souci de continuité et s'appuient sur les aspects des définitions d'origine qui ont prouvé leur efficacité.

Les Orientations devraient être lues en tandem avec les Normes de qualité pour l'évaluation du développement (OCDE, 2010^[4]) et les Principes pour l'évaluation de l'aide au développement (OCDE, 1991^[1]). Les Principes portent essentiellement sur la gestion et les dispositifs institutionnels des systèmes d'évaluation, tandis que les Normes visent à éclairer les processus et les produits d'évaluation. Les présentes Orientations ont vocation à aider les utilisateurs à appliquer les définitions des critères d'évaluation dans la pratique en comprenant mieux les composantes, en fournissant des exemples à l'appui et en analysant les difficultés concrètes généralement rencontrées par les évaluateurs lors de leur mise en œuvre. Encadré 1.1 mentionne les publications à venir qui contribueront à une meilleure application et compréhension des critères.

Au cours des 25 dernières années, beaucoup d'enseignements ont été tirés concernant la manière d'utiliser les critères définis par le CAD, la principale difficulté mise en évidence lors des consultations de 2017-19 étant celle de leur application concrète. Les discussions ont souligné l'importance, non pas seulement de la manière dont les critères sont définis, mais de la façon dont ils devraient être utilisés. Ainsi, les critères sont parfois appliqués mécaniquement et ne sont pas adaptés au contexte. Les présentes Orientations visent par conséquent à aider différents groupes d'utilisateurs à appliquer de manière réfléchie les critères dans la pratique, ce qui contribuera à l'amélioration de la qualité et de l'utilité des évaluations.

Elles emploient donc un langage et formulent des orientations de nature à aider les lecteurs à comprendre et à appliquer les définitions des critères, y compris les deux principes destinés à en guider l'utilisation (OCDE, 2019^[3]). Elles développent ces deux principes et expliquent les définitions de chaque critère, ce qui permet aux utilisateurs d'interpréter et d'appliquer les critères dans le cadre de leurs activités.

Objet des Orientations et mode d'emploi

Le présent document fournit aux utilisateurs un guide succinct sur la manière d'interpréter les critères et de les appliquer de façon réfléchie dans le cadre de leurs activités. Si les critères sont avant tout appliqués dans le domaine de l'évaluation, ils peuvent également l'être pour diverses autres activités comme le suivi et la gestion axée sur les résultats, la conception d'interventions ou la planification stratégique. Les Orientations ont vocation à aider les utilisateurs à réfléchir à certains des défis qui peuvent se poser et à rechercher des solutions. Elles aident le lecteur à appréhender les différentes composantes de chaque

définition et à comprendre comment les critères sont reliés les uns aux autres, quelles sont les difficultés rencontrées dans le cadre de leur application et comment y remédier. Des exemples concrets ont été sélectionnés pour illustrer la façon dont les critères peuvent être appliqués de manière réfléchie dans des contextes réels.

Ces Orientations n'ont pas vocation à servir de manuel, de modèle ou de liste de vérification pour l'évaluation. Le présent guide constitue une source sur laquelle s'appuyer pour élaborer les plans d'évaluation et peut également venir en complément des diverses politiques, normes et orientations institutionnelles sur lesquelles se fondent les évaluateurs.

Encadré 1.1. Autres travaux et publications à venir de l'OCDE sur l'évaluation

Outre la révision des présentes Orientations au regard de l'expérience acquise dans le cadre de l'application des nouvelles définitions, EvalNet et les partenaires s'emploient également à élaborer des orientations spécifiques qui approfondiront certains thèmes, notamment des orientations sur l'utilisation des critères du point de vue de l'égalité des sexes ou sur leur application dans des contextes humanitaires. En collaboration avec EvalNet, l'ALNAP a aussi engagé un processus afin de mettre à jour ses orientations – largement utilisées – sur l'évaluation humanitaire et l'utilisation des critères (Beck, 2016^[5]).

Traduction des définitions des critères et des principes d'utilisation

En collaboration avec ses partenaires clés, et conformément à sa mission consistant à favoriser l'apprentissage et le renforcement des capacités d'évaluation, l'OCDE s'emploie à traduire les définitions des critères dans plusieurs langues, outre les deux langues officielles de l'OCDE (français et anglais). Les traductions en espagnol et en chinois sont achevées et de nouvelles versions en arabe, dari, pachto, portugais, russe, thaïlandais et ourdou sont attendues en 2021.

2^e édition du Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats

Les définitions des cinq critères d'origine ont d'abord été publiées dans le *Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats* (OCDE, 2002^[2]). Une deuxième édition du *Glossaire* est en cours d'élaboration, qui tiendra compte des mises à jour des critères et d'autres termes. Cette 2^e édition (à paraître) accompagnera utilement les présentes Orientations étant donné qu'elle contiendra des définitions succinctes des termes employés tout au long du texte, notamment : « résultats », « réalisations », « impacts », « théorie du changement » et « bénéficiaire ».

De même, les Orientations ne remplacent pas les lignes directrices et normes existantes en matière d'éthique définies par d'autres institutions ou proposées par le CAD de l'OCDE (OCDE, 1991^[1] ; OCDE, 2010^[4]). Elles complètent les normes internationales appliquées par les Nations Unies et d'autres organisations multilatérales, telles que les Normes et règles d'évaluation du Groupe des Nations Unies pour l'évaluation (GNUE), les normes de bonnes pratiques du Groupe de coopération pour l'évaluation (ECG) et les orientations du Réseau pour l'apprentissage actif sur la responsabilité et la performance au sein de l'action humanitaire (Active Learning Network for Accountability and Performance in Humanitarian Action, ALNAP). Entre autres ressources utiles figurent également les sites web des unités d'évaluation d'organismes bilatéraux, des Nations Unies et d'autres organisations multilatérales, le site web « Better Evaluation » et les sites web d'EvalPartners et de ses organisations membres, qui mettent l'accent sur des domaines précis comme l'Initiative internationale pour l'évaluation d'impact (International Initiative for Impact Evaluation, 3ie).

Dans la même veine, les Orientations, de même que les critères eux-mêmes, visent à favoriser l'introspection et la réflexion critique et n'ont pas vocation à être appliqués mécaniquement. Une telle

approche serait incompatible avec les bonnes pratiques et les principes consacrés par le Programme 2030 (voir Chapitre 3 pour plus de détails). Les Orientations ne prescrivent pas d'approches méthodologiques particulières ni d'outil spécifiques, les critères ne formant pas une méthodologie. Des renvois vers d'autres sources qui développent certaines méthodes figurent dans le document à titre de point de départ.

Les Orientations peuvent être considérées comme un pont entre les documents officiels à caractère général (par exemple, les définitions des critères (OCDE, 2019^[3]) et les Normes de qualité de l'OCDE (OCDE, 2010^[4])), qui contiennent des définitions et une terminologie cohérentes et claires, et l'application quotidienne des critères au sein de certaines institutions, d'une manière réfléchie et appropriée, avec les multiples décisions à prendre en fonction du contexte que cela suppose.

Outre le fait qu'elles donnent des exemples de la manière dont les critères ont été appliqués dans le cadre d'évaluations réelles, les Orientations comportent des éléments d'analyse qui expliquent les notions sous-jacentes à chaque définition. Ces éléments ne se veulent pas irréfutables ni ne sont destinés à servir de critères indépendants ou secondaires ; ils visent plutôt à illustrer chaque définition et à permettre d'en appréhender les différentes composantes. Ils aideront les utilisateurs à déterminer quelle est la meilleure interprétation, au regard du contexte, de chaque critère pour telle ou telle évaluation. Au cours de la phase de conception, le dialogue entre les évaluateurs, les commanditaires et les utilisateurs cibles des conclusions de l'évaluation peut également porter sur la manière dont les différents éléments de chaque critère devraient être évalués et sur la question de savoir lesquels sont les plus importants pour tirer des conclusions de portée générale – celles-ci variant selon l'intervention, le contexte et l'utilisation envisagée de l'évaluation.

Le Chapitre 2 s'intéresse au rôle des critères d'évaluation et à leur place dans l'ensemble plus complet de normes et de critères. Le Chapitre 3 formule des orientations sur la manière d'interpréter les critères et de les appliquer de façon réfléchie, et aidera également les évaluateurs à se servir de ces critères dans leur contexte. Le Chapitre 4 présente plus en détail chaque définition, met en lumière les défis communs et fournit des exemples concrets. Eu égard à la priorité stratégique du Programme 2030 visant à « ne laisser personne de côté », ce dernier chapitre offre une vue d'ensemble des principales problématiques en matière d'inclusion qui devraient être examinées en relation avec chaque critère.

Les présentes Orientations sont considérées comme un document de travail, la plupart des exemples provenant d'évaluations effectuées en appliquant les définitions d'origine et le déploiement des nouvelles définitions ainsi que des principes d'utilisation étant toujours en cours. Les Orientations seront utilisées pendant une phase pilote de mise en œuvre initiale, puis révisées après quelques années pour intégrer les enseignements tirés et des exemples à jour. Le Secrétariat cherche tout particulièrement à recueillir d'autres exemples en dehors du cercle d'EvalNet ainsi que des enseignements supplémentaires sur l'utilisation des critères dans le cadre d'évaluations locales et nationales.

Utilisateurs cibles des Orientations

Les Orientations peuvent s'appliquer à divers contextes – local, national et international. Les critères ont une portée délibérément large et ont été conçus pour pouvoir être appliqués à tout un éventail d'interventions et de secteurs. Leur pertinence ne se limite pas à l'évaluation des interventions traditionnelles à l'appui du développement, soit l'aide publique au développement émanant d'États souverains. S'ils sont appliqués de manière réfléchie, ils peuvent également être utiles pour évaluer les interventions relatives au secteur privé, au financement mixte ou à d'autres domaines d'action comme les échanges notamment.

Les Orientations sont destinées principalement aux évaluateurs, aux responsables d'évaluation et aux commanditaires. Elles conviennent tant pour les étudiants ou les jeunes professionnels qui utilisent les critères pour la première fois que pour les professionnels chevronnés qui effectuent des évaluations

complexes. Elles seront également utiles pour guider la conception des interventions, la planification stratégique et la gestion axée sur les résultats. L'amélioration des évaluations dépend en partie de l'intégration efficace des systèmes de suivi, d'évaluation et d'apprentissage dans le programme ou le cycle de l'action publique. Les Orientations aideront les évaluateurs et les partenaires opérationnels à développer dès le départ (c'est-à-dire au cours de la phase de conception) une vision et une terminologie communes de ce qui constitue une intervention réussie, ce qui favorisera l'évaluation des interventions en question, tout en améliorant les résultats globaux.

Références

- Beck, T. (2016), *Evaluating Humanitarian Action using the OECD-DAC Criteria*, ALNAP, [5]
<https://www.alnap.org/help-library/evaluating-humanitarian-action-using-the-oecd-dac-criteria>
 (consulté le 11 janvier 2021).
- OCDE (2019), *Des meilleurs critères pour des meilleures évaluations : Définitions adaptées et principes d'utilisation*, Réseau du CAD sur l'évaluation du développement, Éditions OCDE, Paris, <https://www.oecd.org/fr/cad/evaluation/criteres-adaptees-evaluation-dec-2019.pdf> [3]
 (consulté le 11 janvier 2021).
- OCDE (2010), *Normes de qualité pour l'évaluation du développement*, Lignes directrices et ouvrages de référence du CAD, Éditions OCDE, <https://doi.org/10.1787/9789264083912-fr>. [4]
- OCDE (2002), *L'évaluation et l'efficacité de l'aide n° 6 - Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats (en anglais, français et espagnol)*, Éditions OCDE, Paris, <https://doi.org/10.1787/9789264034921-en-fr>. [2]
- OCDE (1991), *DAC Principles for Evaluation of Development Assistance*, Comité d'aide au développement, Éditions OCDE, Paris, <https://www.oecd.org/development/evaluation/2755284.pdf> (consulté le 11 janvier 2021). [1]

Notes

¹ Le document « De meilleurs critères pour de meilleures évaluations » (OCDE, 2019^[3]) explique dans le détail les modifications apportées aux définitions d'origine ainsi que leur justification.

² Un résumé des conclusions du processus de consultation se trouve sur le site web d'EvalNet : [oe.cd/criteria](https://www.oecd.org/criteria)

2 Les six critères : finalité et rôle dans l'évaluation

Ce chapitre décrit la finalité des six critères d'évaluation et examine le rôle qu'ils jouent dans l'évaluation. Il explique comment les critères peuvent être appréhendés comme un ensemble de prismes complémentaires et la façon dont chacun peut apporter une perspective différente sur l'intervention et ses résultats. Il présente également différents types d'interventions dont l'intérêt ou le mérite peut être évalué en se servant des critères comme base d'appréciation. La dernière partie analyse les liens entre les critères et d'autres normes et critères d'évaluation.

Qu'est ce qu'un critère d'évaluation ?

Un critère est une norme ou un principe utilisé(e) dans l'évaluation comme base de jugement évaluatif.

Chacun des six critères est résumé sous la forme d'une question générale qui en illustre le champ d'analyse. Chaque critère représente un élément d'examen important :

- Pertinence : L'intervention¹ répond-elle au problème ?
- Cohérence : L'intervention s'accorde-t-elle avec les autres interventions menées ?
- Efficacité : L'intervention atteint-elle ses objectifs ?
- Efficience : Les ressources sont-elles utilisées de manière optimale ?
- Impact : Quelle différence l'intervention fait-elle ?
- Viabilité/Durabilité : Les bénéfices seront-ils durables ?

La finalité des critères d'évaluation est de contribuer à une évaluation cohérente et de qualité menée conformément à un cadre commun. Les critères offrent un cadre normatif qui permet d'évaluer une intervention donnée. Ils peuvent également être utilisés pour d'autres processus que celui de l'évaluation, notamment pour la définition de cadres et d'indicateurs de suivi et de gestion axée sur les résultats, l'approbation de financements, la planification stratégique et la conception d'interventions, en particulier en vue d'améliorer celles qui seront menées à l'avenir. Collectivement, il est utile de disposer de critères définis d'un commun accord qui soient appliqués de manière identique quelle que soit l'intervention. Les critères offrent également une terminologie cohérente pour l'ensemble du domaine du développement, favorisant ainsi la normalisation, les comparaisons et l'apprentissage d'une intervention à l'autre.

Les critères devraient être considérés comme une série de prismes à travers lesquels une intervention peut être appréhendée et analysée. Ils apportent des perspectives complémentaires et permettent d'avoir une vue globale de l'intervention. Ils encouragent l'approfondissement de la réflexion sur la nature d'une intervention, sa mise en œuvre, son processus et ses résultats. Collectivement, ils décrivent les caractéristiques attendues de toutes les interventions, formulent des hypothèses explicites et fixent des normes pour les interventions, à savoir : être adaptées au problème à résoudre, être cohérentes avec les autres interventions, produire des résultats de manière efficiente et se traduire par des effets positifs pérennes sur le développement durable. Les critères d'évaluation offrent également un cadre largement reconnu pour définir une approche de l'évaluation qui soit complète et systématique ainsi qu'une terminologie commune à utiliser dès le début du processus d'évaluation.

Les critères sont tous reliés entre eux et peuvent être considérés comme un ensemble complémentaire dans lequel chaque critère apporte une perspective différente et unique pour appréhender l'intervention. Même si la définition de chaque critère incarne une notion particulière (illustrée par la question générale qui lui est associée), ces notions sont à de multiples égards liées. Les interventions de développement ont tendance à être multidimensionnelles ; aussi, l'utilisation combinée des différents critères peut aider l'évaluateur à examiner l'intervention dans son ensemble.

Les critères d'évaluation ne forment pas une méthodologie, ni ne constituent des objectifs qu'une intervention tenterait d'atteindre. Ils constituent plutôt des incitations à se poser les bonnes questions lors de l'évaluation d'une intervention. Parce que chaque intervention (l'objet de l'évaluation ou le sujet évalué²) est différente, le processus d'évaluation doit être souple et l'utilisation des critères d'évaluation doit être réfléchi et adaptée à la finalité de l'évaluation et aux utilisateurs. Conformément aux normes de qualité des évaluations, les conclusions et les recommandations relatives aux progrès et aux performances devraient se fonder sur des données pertinentes et crédibles. La logique, la crédibilité et l'interprétation des données devraient être claires et explicites et se conformer à la théorie du changement ou à la logique de l'intervention.

Quels types d'interventions peuvent être évalués à l'aune de ces critères ?

Le terme « intervention » est employé tout au long du document pour désigner le thème ou l'objet de l'évaluation. Il englobe tous les types d'activités qui peuvent être évalués à l'aune de ces critères. Ces activités peuvent être menées à l'échelon international ou national et peuvent viser à favoriser le développement durable ou à satisfaire à des objectifs humanitaires.

Il peut s'agir d'un projet, d'un programme, d'une politique, d'une stratégie, d'un domaine thématique, d'un programme d'assistance, d'une activité de conseil sur l'action publique, d'une initiative institutionnelle, d'un mécanisme de financement, d'un instrument ou de toute autre activité. Sont compris les interventions de développement, l'aide humanitaire, l'appui à la consolidation de la paix, les travaux normatifs et d'autres activités de coopération internationale, ainsi que les activités des acteurs du secteur privé et des administrations nationales dans le contexte du pays. À titre d'illustration, les critères ont déjà été utilisés pour évaluer des thèmes allant de l'efficacité des programmes nationaux sur l'alimentation scolaire à la cohérence du soutien international apporté par différents acteurs dans une région touchée par un conflit. Ils servent également à évaluer des politiques ou des stratégies (voir, par exemple, Encadré 2.1) et sont aussi utilisés pour évaluer des séries d'interventions – par exemple, un ensemble de projets à l'appui du secteur de l'éducation ou une intervention unique de grande ampleur étalée sur de nombreuses années et impliquant de multiples partenaires (soutien budgétaire général, etc.).

Si le terme « intervention » est employé dans ce document à des fins de facilité de compréhension, les évaluateurs devraient veiller à définir clairement le thème de l'évaluation – l'intervention analysée – ainsi que le périmètre couvert à un stade précoce du processus. L'emploi d'autres termes plus précis liés à l'intervention en question – par exemple, « projet » ou « politique » – sera probablement plus utile et intelligible pour les partenaires.

Encadré 2.1. Adapter les critères à l'intervention : évaluation du plan d'action du BMZ pour l'inclusion des personnes handicapées

L'évaluation du plan d'action du ministère fédéral allemand de la Coopération économique et du développement (BMZ) pour l'inclusion des personnes handicapées offre un exemple intéressant de la manière dont les critères peuvent être adaptés à l'intervention évaluée (soit, dans le cas présent, un plan d'action stratégique) et sont rattachés aux questions d'évaluation. Cette évaluation visait à examiner dans quelle mesure le plan d'action avait réussi à promouvoir l'intégration systématique du principe d'inclusion dans les activités allemandes de coopération pour le développement. Elle reposait sur une série de questions d'évaluation d'ordre général (voir ci-après), tirées du plan d'action lui-même. Chacune de ces questions était ventilée en plusieurs questions plus précises par critère :

1. Dans quelle mesure le BMZ montre-t-il l'exemple en matière d'inclusion des personnes handicapées dans sa propre organisation interne ?

- Pertinence : Dans quelle mesure les domaines d'action et mesures sélectionnés correspondent-ils aux dispositions de la Convention des Nations Unies relative aux droits des personnes handicapées (CDPH) ?
- Efficacité : Dans quelle mesure le BMZ parvient-il à établir des structures et des pratiques inclusives ?

2. Dans quelle mesure le plan d'action contribue-t-il à favoriser l'inclusion des personnes handicapées dans les pays partenaires de la coopération allemande pour le développement ?

- Pertinence : Dans quelle mesure les domaines d'action et mesures sélectionnés correspondent-ils aux dispositions de la CDPH ?

- Efficacité : Dans quelle mesure l'inclusion des personnes handicapées a-t-elle été intégrée avec plus d'efficacité dans les mesures prises dans les pays partenaires de la coopération allemande pour le développement ?

3. Dans quelle mesure le BMZ joue-t-il, aux niveaux national, régional et international, le rôle de défenseur et de partenaire des droits des personnes handicapées dans le cadre de la coopération pour le développement ?

- Pertinence : Dans quelle mesure les domaines d'action et mesures sélectionnés correspondent-ils aux dispositions de la CDPH ?
- Efficacité : Dans quelle mesure le BMZ est-il parvenu à rallier les acteurs bilatéraux et multilatéraux à la cause de l'inclusion ?

4. Comment l'élaboration et la mise en œuvre du plan d'action pour l'inclusion des personnes handicapées ont-elles été gérées ?

- Efficacité : Quels mécanismes et structures de gestion (y compris les rôles joués par les divers groupes de parties prenantes dans le processus de gestion) ont déterminé l'élaboration et la mise en œuvre du plan d'action ?
- Efficacité : Dans quelle mesure les mesures ont-elles été mises en œuvre comme prévu ?

5. Comment devrions-nous noter les bénéfices du plan d'action pour l'inclusion des personnes handicapées en termes d'ampleur de l'impact et d'effet de levier en tant qu'instrument de gouvernance ?

- Impact : Dans quelle mesure le plan d'action a-t-il eu un impact à grande échelle ?
- Impact : Dans quelle mesure le plan d'action a-t-il eu un effet de levier ?

Source : Schwedersky, Ahrens et Steckhan (2017^[1]), *Evaluation of the BMZ Action Plan for the Inclusion of Persons with Disabilities*, <https://d-nb.info/1186644206/34>

Quelle est la place des critères dans le domaine de l'évaluation en général ?

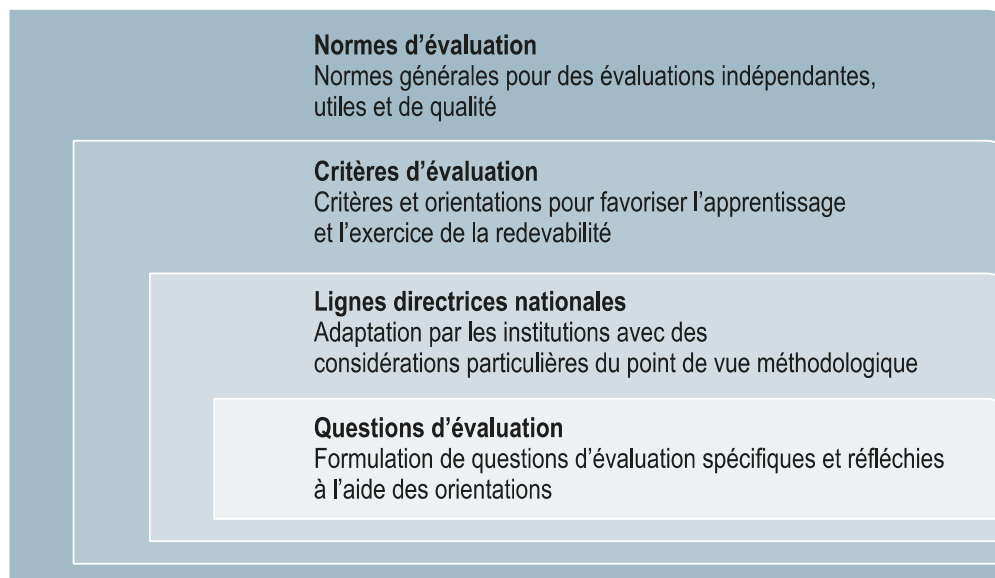
Parce que chaque intervention (l'entité évaluée) est différente, le processus d'évaluation doit être souple et l'utilisation des critères d'évaluation doit être réfléchie et adaptée à la finalité de l'évaluation et aux utilisateurs. Les critères d'évaluation n'ont donc pas vocation à poser un ensemble de règles obligatoires.

Ils n'existent pas isolément. Pour en faire le meilleur usage, il importe de comprendre comment ils se placent par rapport aux autres normes et règles, méthodologies et orientations institutionnelles.

- Au macro-niveau se situent les principes fondamentaux d'évaluation tels que l'impartialité et l'indépendance, la crédibilité, l'utilité et la participation (OCDE, 1991^[2]), ainsi que les *Normes de qualité pour l'évaluation du développement* (OCDE, 2010^[3]). Les normes d'éthique dans le domaine de la recherche et de la collecte de données relèvent également de ce niveau, qui regroupe donc les normes générales de bonnes pratiques auxquelles toute évaluation devrait se conformer, tout au long du processus.
- Au méso-niveau se situent les critères d'évaluation, qui fournissent les prismes à travers lesquels une intervention est analysée et appréhendée.
- Le niveau institutionnel désigne le niveau auquel chaque organisation adapte les critères et les traduit en lignes directrices et politiques internes (par exemple, en mettant l'accent sur certains critères, en exigeant des notations, en ajoutant des critères supplémentaires) qui reflètent leurs propres mandats et priorités.

- Au micro-niveau, c'est-à-dire dans le contexte de chaque évaluation, se situent les décisions relatives aux questions d'évaluation et aux méthodologies appliquées pour chaque évaluation, ainsi que la manière dont les critères sont appliqués dans ce contexte et pour le sujet évalué.

Graphique 2.1. Comment se placent les critères par rapport aux autres normes et critères ?



Références

- OCDE (2010), *Normes de qualité pour l'évaluation du développement*, Lignes directrices et ouvrages de référence du CAD, Éditions OCDE, Paris, <https://doi.org/10.1787/9789264083912-fr>. [3]
- OCDE (1991), *Principles for Evaluation of Development Assistance*, Comité d'aide au développement, Éditions OCDE, Paris, <https://www.oecd.org/development/evaluation/2755284.pdf> (consulté le 11 janvier 2021). [2]
- Schwedersky, T., L. Ahrens et H. Steckhan (2017), *Evaluation of the BMZ Action Plan for the Inclusion of Persons with Disabilities*, DEval, <https://d-nb.info/1186644206/34> (consulté le 11 janvier 2021). [1]

Notes

¹ Le terme « intervention » est employé ici et tout au long du document pour désigner l'objet de l'évaluation (ce qui est évalué). Se reporter au Chapitre 3, section « Adapter les critères à la finalité de l'évaluation », pour un examen plus approfondi en la matière

² L'expression « sujet évalué » fait référence à l'objet de l'évaluation – ce qui est évalué.

3

Utiliser les critères d'évaluation dans la pratique

Le présent chapitre traite des aspects pratiques liés à la manière dont les critères devraient être utilisés dans le cadre de la conception et de la mise en œuvre des évaluations. Il examine les deux principes fondamentaux qui devraient guider l'application des critères. Après avoir analysé différentes approches susceptibles de favoriser une application réfléchie des critères, il s'intéresse à la façon dont les critères peuvent être appliqués dans différents contextes institutionnels et avec des priorités stratégiques, des méthodes de travail et des cultures différentes. Il décrit sommairement comment les critères peuvent inciter les évaluateurs à examiner les différences en termes d'expérience et d'impact en appliquant un prisme sexospécifique. Enfin, il examine comment les critères peuvent aider les évaluateurs et les responsables d'évaluation à travailler d'une manière qui contribue à la réalisation des Objectifs de développement durable (ODD) et, plus globalement, au Programme de développement durable à l'horizon 2030. Divers exemples concrets ponctuent le chapitre, qui montrent comment les critères ont été appliqués dans le cadre de diverses évaluations.

Pour aider les évaluateurs et les personnes chargées de la conception ou de la gestion des interventions à mettre au point des évaluations utiles et adaptées aux différents contextes et parties prenantes, les deux principes suivants ont été élaborés pour guider l'utilisation des critères. Pour éviter que les critères ne soient appliqués mécaniquement – aux dépens de la réflexion critique, de la créativité et de l'appropriation par les participants –, ces principes devraient être pris en compte dès lors que les critères sont utilisés (OCDE, 2019^[1]) :

- **Principe n°1** : Les critères devraient être appliqués de façon réfléchie pour renforcer la qualité et l'utilité des évaluations.
- **Principe n°2** : L'utilisation des critères dépend de la finalité de l'évaluation.

La section qui suit développe ces deux principes et présente d'autres notions clés liées à l'utilisation des critères, notamment la manière de les adapter à des contextes particuliers, la manière de les examiner à différents moments dans le temps et la manière dont ils sont reliés les uns aux autres.

Une application réfléchie des critères

Le Principe n°1 souligne que les critères devraient être appliqués de façon réfléchie. Dans la pratique, cela signifie mener une réflexion critique sur les critères qui sont le plus utiles pour contribuer à une évaluation utile et de qualité, présentant un intérêt pour les utilisateurs auxquels elle est destinée. L'Encadré 3.1 illustre la manière dont les critères peuvent être appliqués de façon réfléchie lors de l'évaluation d'une intervention.

La prise en compte des six aspects suivants et des questions qui leur sont associées aidera les évaluateurs à utiliser les critères de façon réfléchie :

- **Contexte** : Quel est le contexte de l'intervention et comment les critères peuvent-ils être appréhendés dans le contexte de l'évaluation, de l'intervention et des parties prenantes concernées ?
- **Finalité** : Quel est le but que l'évaluation tente d'atteindre et quelles questions sont les plus utiles pour poursuivre et atteindre ce but ?
- **Rôles et dynamiques de pouvoir** : Qui sont les parties prenantes, quels sont leurs besoins et intérêts respectifs ? Quelles dynamiques de pouvoir sont à l'œuvre ? Qui doit être associé à la prise de décisions sur les critères à appliquer et comment appréhender ces critères dans le contexte spécifique de l'évaluation et de l'intervention ? Peuvent ici être posées des questions sur l'appropriation et sur les personnes qui décident de l'objet et des priorités de l'évaluation.
- **Intervention (sujet évalué)** : Quel type d'intervention est évalué (un projet, une politique, une stratégie, un secteur) ? Quelle est sa portée et sa nature ? Dans quelle mesure ses résultats attendus sont-ils directs ou indirects ? Quelle théorie des systèmes complexes entre en jeu ?
- **Évaluabilité** : Existe-t-il des contraintes en termes d'accès, de ressources et de données (y compris les données ventilées) qui ont une incidence sur l'évaluation, et en quoi ces contraintes influent-elles sur les critères ?
- **Temporalité** : À quel stade du cycle de vie de l'intervention l'évaluation sera-t-elle effectuée ? Le contexte dans lequel s'inscrit l'intervention a-t-il évolué avec le temps et, dans l'affirmative, de quelle manière ? Ces évolutions devraient-elles être prises en compte pendant l'évaluation ? Le moment choisi pour l'évaluation influera sur l'utilisation des critères ainsi que sur les sources de données.

Encadré 3.1. Utilisation réfléchie des critères : évaluation de la dotation de la Norvège en faveur de la société civile dans les pays en développement

Cette évaluation visait à examiner la dotation de la Norvège en faveur de la société civile, qui a pour objet de renforcer la société civile dans les pays en développement et, partant, de contribuer à une société civile plus forte, dotée de la capacité et des moyens nécessaires pour promouvoir la démocratisation, améliorer les droits humains et réduire la pauvreté.

Dans le cadre de cette évaluation, les critères ont été appliqués de façon « réfléchie » et souple, et interprétés d'une manière particulière tenant compte de la nature de l'intervention (partenariats avec des organisations de la société civile et renforcement de leurs capacités) ainsi que des priorités norvégiennes et du contexte. L'évaluation s'appuyait sur les critères suivants, qui avaient été adaptés à certains égards :

- impact – défini au regard des principales réalisations, par exemple la démocratie ou les possibilités de revenus entraînant une réduction de la pauvreté
- pertinence – aucun changement
- efficacité – définie au regard des objectifs spécifiques du partenariat (c'est-à-dire prestation de services, sensibilisation et renforcement des capacités). Englobait certaines réalisations à moyen terme, comme la création d'un espace dévolu à la société civile
- viabilité/durabilité – aucun changement.

Un autre critère, celui de la « valeur ajoutée », était également utilisé, défini comme la compétence professionnelle, la compétence organisationnelle et financière et la compétence relationnelle (réseaux). Ce critère se rapportait à des aspects qui, dans d'autres contextes, pourraient relever des critères d'efficacité et d'efficience.

Source : Helle et al. (2018^[2]), *From Donors to Partners? Evaluation of Norwegian Support to Strengthen Civil Society in Developing Countries through Norwegian Civil Society Organisations*, https://www.norad.no/globalassets/filer-2017/evaluating/1.18-from-donor-to-partners/1.18-from-donors-to-partners_main-report.pdf

Adapter les critères à la finalité de l'évaluation

Lorsqu'il s'agit de décider de la manière dont les critères devraient être utilisés, le plus important est de les relier à la finalité de l'évaluation et à son contexte, puis de les articuler, ainsi que les questions d'évaluation, autour de cette finalité. L'Encadré 3.2 fournit deux exemples de la manière dont la finalité d'une évaluation peut être définie.

Les critères n'ont pas vocation à être appliqués de manière uniformisée ni figée à chaque intervention, ni ne sont censés servir de liste de vérification. En réalité, ils devraient être interprétés ou appréhendés avec circonspection, au regard de l'intervention évaluée. Une telle démarche favorise la souplesse et l'adaptation des critères à chaque évaluation. Il faudrait clairement indiquer sur quelles notions sous-jacentes aux critères l'évaluation s'appuiera en particulier, et pourquoi.

La finalité de l'évaluation devrait être soigneusement et clairement définie. Les parties prenantes qui participent à l'évaluation devraient être associées à ce stade afin de s'assurer qu'elles comprennent le but de l'évaluation et l'utilisation qui en sera faite.

Parmi les principales questions à examiner pour déterminer la finalité de l'évaluation figurent les suivantes :

1. Quel est le besoin d'évaluation, qui est le public cible et quel usage fera-t-il des constatations de l'évaluation ?
2. Quelles actions sont réalisables compte tenu des caractéristiques et du contexte de l'intervention ?
3. Quel degré de certitude est nécessaire pour pouvoir répondre aux questions clés ?
4. À quel moment les informations sont-elles nécessaires ?
5. Que sait-on d'ores et déjà de l'intervention et de ses résultats ? Qui est détenteur de ces informations et comment sont-elles mises à profit ?

Les normes de qualité et en matière d'éthique devraient éclairer la réflexion subséquente sur la méthodologie, la conception, la mise en œuvre et la gestion du processus d'évaluation.

Pour adapter les critères à une évaluation donnée, il importe également d'analyser en profondeur les liens de causalité et d'évaluer dans quelle mesure l'évaluation parviendra à attribuer des effets à l'intervention évaluée. Ces considérations peuvent aider à gérer les attentes des parties prenantes et à déterminer les critères qui devront être examinés en profondeur. Une telle démarche est particulièrement importante pour les critères d'efficacité et d'impact (voir les deux sections qui suivent), mais est aussi indirectement applicable – et peut servir – à d'autres critères. Par exemple, les bénéfices réels (ou projetés) attribués à l'intervention au titre de l'évaluation de l'efficacité et de l'impact pourraient également servir à l'évaluation de l'efficacité et de la viabilité/durabilité. L'analyse de l'évaluabilité est utile pour s'assurer que la finalité de l'évaluation est parfaitement comprise¹. Le Chapitre 4 contient des explications supplémentaires sur la manière d'interpréter chaque critère.

Encadré 3.2. Définir la finalité de l'évaluation : exemples de l'Ouganda, du Kenya et de la Suède

Évaluation par l'Agence danoise de développement international (Danida) des programmes relatifs à l'eau, à l'assainissement et à l'environnement en Ouganda

Le Département de l'évaluation du ministère danois des Affaires étrangères a commandé une évaluation indépendante des initiatives danoises menées sur la période 1990-2017 en vue de développer les secteurs de l'eau, de l'assainissement et de l'environnement en Ouganda.

Les objectifs de l'évaluation étaient les suivants :

- recenser les résultats et les réalisations dans les sous-secteurs ;
- analyser la « valeur ajoutée » du soutien danois à ces sous-secteurs ;
- dégager des enseignements.

À ces fins, l'évaluation a mis l'accent sur l'évaluation de l'efficacité à deux niveaux : 1) sectoriel et transversal et 2) sous-sectoriel. Elle a également apprécié l'intérêt des initiatives et examiné dans quelle mesure les résultats obtenus étaient durables.

Évaluation conjointe du Haut-commissariat des Nations Unies pour les réfugiés (HCR) et du Danemark

Le but de l'évaluation menée conjointement par le HCR et le Danemark concernant l'action en faveur des réfugiés dans le camp de Kalobeyei, au Kenya, et alentour est défini comme suit :

L'évaluation a pour principal objectif de contribuer à l'apprentissage concernant le modèle de camp intégré mis en œuvre à Kalobeyei et alentour. En recensant les enseignements tirés de la mise en œuvre d'une initiative concrète visant à relier aide humanitaire et aide au développement à long terme, l'intention est d'apporter des informations sur les possibilités et les défis liés à la conception et à la mise en œuvre d'un modèle de solution intégrée.

Évaluation du soutien apporté par l'Agence suédoise de coopération pour le développement international (Asdi) en faveur de la consolidation de la paix dans les situations de conflit et d'après-conflit

Le rapport d'évaluation donne un exemple utile de la manière dont l'évaluation reflète le but visé. Ainsi qu'il l'explique :

L'évaluation a pour but ou utilité envisagée de systématiser les enseignements tirés de la pratique de la consolidation de la paix. Elle servira d'intrant dans le processus de conceptualisation et de formulation de l'approche que l'Asdi appliquera dans ce domaine. Cette approche influera à son tour sur la planification stratégique et la conception des activités futures de soutien de l'Asdi dans les contextes touchés par un conflit. En outre, l'évaluation est censée contribuer à une meilleure compréhension de la consolidation de la paix en tant que concept et pratique.

L'évaluation visait spécifiquement à évaluer comment l'Asdi aborde la problématique de la consolidation de la paix au niveau stratégique dans des contextes différents. À cette fin, elle a accordé une attention particulière à quatre critères :

- la pertinence des activités de consolidation de la paix menées par l'Asdi par rapport aux besoins du contexte et des bénéficiaires, et ses priorités stratégiques ;
- l'efficacité de ses activités de consolidation de la paix en termes de contribution aux objectifs globaux dans ce domaine et son aptitude à offrir un cadre propice aux activités de consolidation de la paix menées par ses partenaires ;
- l'impact de ses activités de consolidation de la paix ;
- la pérennité de ses activités de consolidation de la paix.

Pour l'ensemble de ces quatre critères, un accent particulier a été mis sur les résultats dans les domaines de l'égalité entre les sexes, de l'autonomisation et des droits des femmes, et de l'inclusion des groupes marginalisés et des minorités ethniques. Les évaluations relatives aux quatre études de cas ont porté principalement sur des dimensions de la marginalisation liées à la consolidation de la paix et abordées dans la littérature ainsi qu'au cours des entretiens. La marginalisation des groupes a été principalement considérée sous l'angle de l'appartenance ethnique ou comme une conséquence de l'appartenance à une minorité, les facteurs ethniques constituant la principale cause de conflit dans les quatre cas étudiés.

Source : Danida (2019^[3]), *Evaluation of Water, Sanitation and Environment Programmes in Uganda (1990-2017)*, <http://www.oecd.org/derec/denmark/denmark-1990-2017-wash-environment-uganda.pdf> ;

ADE (2019^[4]), *Joint Evaluation of the Integrated Solutions Model in and around Kalobeyei, Turkana, Kenya*, https://um.dk/en/danida-en/results/eval/eval_reports/publicationdisplaypage/?publicationid=dd54bef1-1152-468c-974b-c171fbc2452d ;

Bryld (2019^[5]), *Evaluation of Sida's Support to Peacebuilding in Conflict and Post-Conflict Contexts: Somalia Country Report*, https://publikationer.sida.se/contentassets/1396a7eb4f934e6b88e491e665cf57c1/eva2019_5_62214en.pdf

Comprendre la dimension temporelle

Les critères peuvent être appliqués à différents moments dans le temps. Chacun d'entre eux peut être utilisé pour évaluer une intervention avant, pendant ou après sa mise en œuvre. De même, ils peuvent être évalués à différents moments du cycle de vie de l'intervention. Toutefois, l'interprétation des critères et des sources de données peut varier selon le moment choisi. Ainsi, avant une intervention, l'efficacité et la viabilité/durabilité sont de simples projections, tandis qu'après l'intervention, des données plus nombreuses sont disponibles, qui permettent de tirer des conclusions plus robustes.

Les critères – et les questions d'évaluation qui leur sont associées – devraient donc tenir compte des deux principaux aspects temporels suivants : 1) le moment, dans le cycle de vie de l'intervention, où s'effectue l'évaluation et 2) le stade où se trouve l'intervention ou le moment dans la chaîne de résultats sur lequel l'évaluation se concentrera.

Le cadre conceptuel de chaque critère reste le même, que l'évaluation soit effectuée avant, pendant ou après une intervention. En revanche, les données et les autres éléments factuels disponibles pour *évaluer* les critères et les méthodes utilisées changent. Les évaluateurs devraient garder ces différences à l'esprit lorsqu'ils examinent les constatations (potentielles) avec les parties prenantes de l'évaluation, car elles peuvent influencer sur la manière dont l'utilité et la crédibilité de l'évaluation sont perçues. Ainsi, une évaluation *ex ante* de la viabilité/durabilité (avant le démarrage d'une intervention) pourra évaluer la probabilité que les bénéficiaires de l'intervention perdurent dans le temps, en examinant la conception de l'intervention et les données disponibles attestant la validité des hypothèses relatives à la pérennité des bénéficiaires attendus. Une fois l'intervention achevée, l'évaluation de la viabilité visera à déterminer si les bénéficiaires ont effectivement été pérennes, cette fois-ci en s'appuyant sur des données et d'autres éléments factuels tirés des bénéficiaires réellement obtenus.

Les évaluations rétrospectives devront prendre en compte le contexte et les données disponibles au moment en question pour opérer des jugements sur la base d'attentes raisonnables quant aux actions qui auraient pu ou dû être menées. Il serait injuste de juger les actions des concepteurs de programmes achevés au regard des informations disponibles au moment de l'évaluation et non de celles alors connues des intéressés. Cela étant, il a été observé dans le cadre de nombreuses évaluations que les informations disponibles n'avaient pas été pleinement utilisées (là où elles auraient pu l'être), ce qui aurait pu accroître la pertinence de l'intervention. Par exemple, les populations locales n'avaient pas été suffisamment consultées ni associées au stade de la conception de l'intervention. Cette lacune, dont on aurait raisonnablement pu s'attendre à ce qu'elle soit évitée, devrait donc être signalée dans une évaluation.

Utiliser les critères pour améliorer les questions d'évaluation

La formulation de questions d'évaluation adéquates constitue une partie essentielle du processus d'évaluation et l'utilisation des six critères interagit avec la prise de décisions sur ces questions, à laquelle elle contribue². Le processus débute par une réflexion sur la finalité de l'évaluation, la manière dont elle sera mise à profit et par qui. Un dialogue efficace avec les parties prenantes, à travers un processus participatif bien conçu, peut aider les évaluateurs et les responsables à comprendre comment ces parties mettront à profit l'évaluation. Ce dialogue devrait être complété par une compréhension profonde de l'intervention et du contexte dans lequel elle s'inscrit, ainsi que de ses objectifs et de la théorie du changement appliquée. Une phase initiale peut servir à étudier ces aspects et à formuler (en petit nombre) des questions clés auxquelles l'évaluation devra répondre.

Après cette phase initiale, les critères d'évaluation permettent d'opérer des vérifications sous plusieurs angles afin de repérer les éventuelles omissions. Il est ainsi possible de développer et d'affiner les questions – un aspect essentiel pour la conception de l'évaluation. Une telle démarche permet de systématiser le processus et garantit l'exhaustivité de l'évaluation. C'est un élément essentiel d'une « meilleure évaluation ». Il est fortement recommandé d'adopter une approche globale claire et cohérente pour cette phase d'élaboration des questions d'évaluation.

Il se peut que l'institution commanditaire ait pris la décision d'évaluer les interventions selon certains critères. Dans ce cas, l'évaluation devrait être menée conformément aux orientations de l'institution et les critères devraient être appliqués de façon réfléchie, à des fins de cohérence.

Le manuel d'évaluation de l'Asdi en est une illustration, qui fournit des exemples de questions types en relation avec chaque critère (Molund et Schill, 2004^[6]). Les orientations et les notes techniques élaborées

à l'intention des responsables d'évaluation et des évaluateurs du Programme alimentaire mondial constituent un autre exemple de la manière dont les institutions interprètent les critères et en appréhendent les composantes (PAM, 2016^[7]).

Liens entre les critères

Les critères constituent de multiples prismes à travers lesquels une intervention et ses résultats peuvent être considérés. Ils sont liés entre eux dans le sens où les notions qui les sous-tendent permettent d'analyser des dimensions complémentaires du processus d'obtention des résultats. Par exemple, l'efficacité et l'impact examinent des niveaux différents de la chaîne de résultats, selon la manière dont les objectifs de l'intervention ont été définis. Il existe donc une corrélation entre ces deux critères tout au long de la chaîne de causalité.

Les critères sont souvent interdépendants. Ainsi, une intervention ne présentant pas de pertinence au regard des priorités des bénéficiaires a peu de chances de produire l'impact attendu (à moins que celui-ci puisse être assuré par d'autres voies inédites). Une intervention mal mise en œuvre (moins efficace) risque également de se traduire par des effets escomptés. À l'inverse, il peut aussi arriver qu'une intervention hautement pertinente soit inefficace, ou qu'une intervention très cohérente soit particulièrement inefficace en raison de coûts de transaction accrus. Les évaluateurs devraient étudier et analyser les liens et les synergies entre les différents critères, en s'intéressant notamment à la question de savoir s'ils sont liés par un lien de causalité et dans ce cas, de quelle manière.

La plupart des évaluations tirent des conclusions à partir des constatations dressées pour chaque critère ainsi qu'une conclusion générale, et notent parfois la performance à l'aide d'une valeur numérique. Pour tirer des conclusions sur l'intervention, et selon le but de l'évaluation, les évaluateurs devraient examiner la situation dans son ensemble et réfléchir à la manière de pondérer comme il se doit chacun des critères appliqués. À cet égard, certaines institutions définissent parfois un critère dominant (éliminatoire). Si les résultats ne sont pas satisfaisants au regard de ce critère, aussi bons les résultats obtenus au niveau des autres critères soient-ils, l'intervention est considérée comme infructueuse (ou, s'il s'agit d'une évaluation *ex ante*, elle n'est pas financée).

On trouvera un examen de liens plus spécifiques entre chaque critère au Chapitre 4.

Sélectionner les critères à utiliser

Comme indiqué précédemment, la finalité, les priorités, la portée et le contexte de l'intervention et de l'évaluation détermineront l'importance relative accordée à chaque critère. Les évaluateurs devraient considérer la valeur ajoutée relative qu'apportera chaque critère, ce qui suppose de répondre à deux questions fondamentales : Ce critère est-il un élément d'appréciation important pour l'évaluation ? Est-il possible de répondre aux questions relatives à ce critère ?

Si les utilisateurs peuvent être tentés d'appliquer automatiquement l'ensemble des six critères, quel que soit le contexte, la meilleure approche (c'est-à-dire, celle qui est cohérente avec l'intention initiale des critères et qui permettra de mener une évaluation de qualité supérieure) consiste à sélectionner délibérément les critères à appliquer et à les utiliser d'une manière adaptée à l'évaluation et aux questions auxquelles cette dernière vise à répondre.

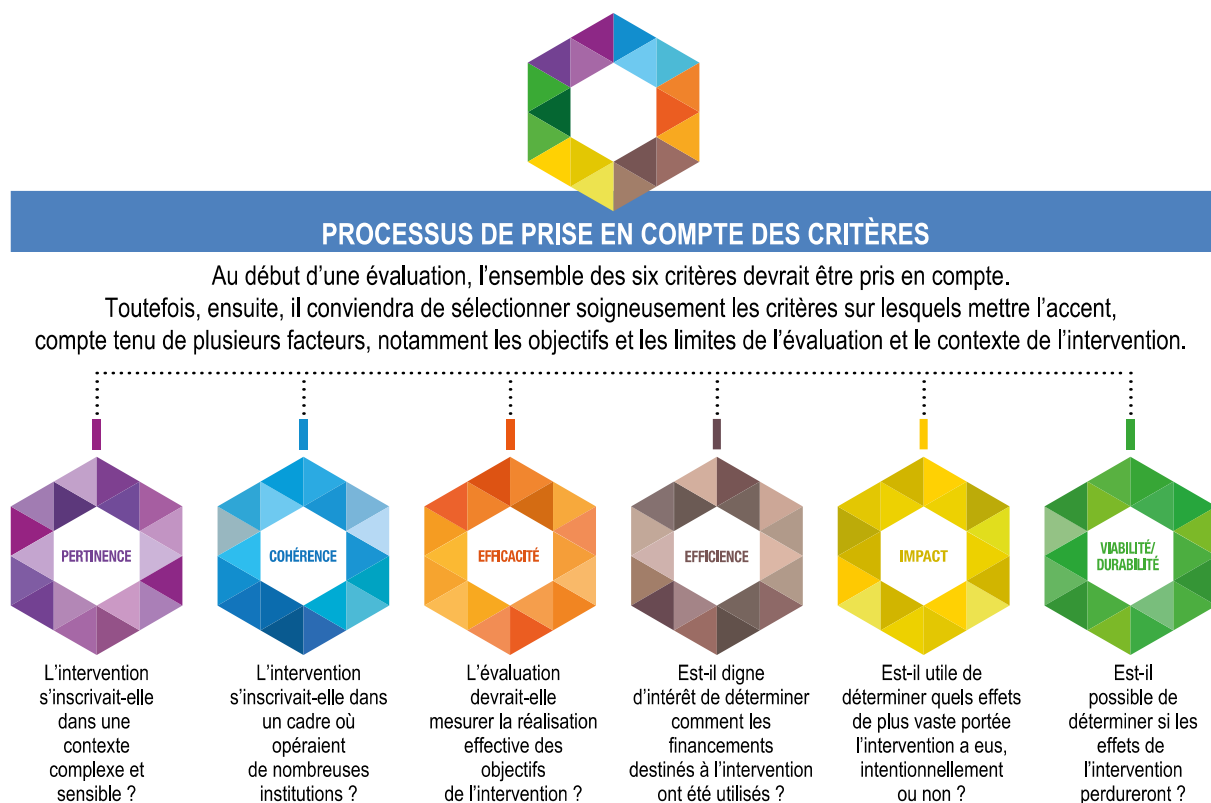
Pour y parvenir, il convient de se poser des questions du type :

- Si une seule question pouvait être posée concernant cette intervention, quelle serait-elle ?

- Quelles questions trouvent leur meilleure réponse en menant une évaluation et lesquelles en recourant à d'autres moyens (par exemple, un projet de recherche, une synthèse de données, un exercice de suivi ou un processus d'apprentissage itératif) ?
- Les données disponibles sont-elles suffisantes pour apporter une réponse satisfaisante à cette question ? À défaut, des données de meilleure qualité ou en plus grand nombre seront-elles disponibles par la suite ?
- Qui a apporté sa contribution pour élaborer la liste de questions ? Est-ce qu'il manque des angles de vue importants ?
- Disposons-nous de suffisamment de temps et de ressources pour prendre en compte de manière adéquate l'ensemble des critères qui présentent un intérêt, ou obtiendrons-nous des informations plus utiles en axant l'analyse sur certains critères seulement ?

Lorsqu'il s'agit d'utiliser les critères, il importe de trouver un équilibre entre flexibilité (éviter toute application mécanique des critères dans leur ensemble) et « picorage » (choisir uniquement les critères les plus faciles ou ceux qui sont les plus susceptibles de déboucher sur des résultats positifs). En particulier, il ne faut pas reculer devant la tâche parfois difficile de répondre à des questions essentielles sur l'impact et la cohérence. Le Graphique 3.1 ci-après présente certains aspects à prendre en compte ; il convient de noter qu'il s'agit d'exemples et non d'une liste exhaustive.

Graphique 3.1. Réfléchir à la manière de prendre en compte les critères pour une évaluation donnée



S'agissant de la conception et de la mise en œuvre de l'évaluation, une bonne connaissance des parties prenantes concernées – par l'intervention mais aussi par l'évaluation – peut permettre de repérer les tensions potentielles entre leurs intérêts et leurs priorités. Ainsi, l'intérêt des bénéficiaires pourra porter

davantage sur l'efficacité (par exemple, le fait de savoir si la participation à un programme de traitement de la malnutrition permet d'améliorer la santé de leurs enfants), tandis que les organismes d'exécution se soucieront plutôt de l'efficacité (dans l'optique de proposer le traitement à un plus grand nombre de familles, par exemple). Dans la plupart des cas, une évaluation ne peut seule répondre à toutes les questions potentiellement intéressantes et des choix doivent être opérés. Pour accroître la probabilité que les questions non couvertes par l'évaluation le soient dans le cadre d'un autre processus, une bonne pratique consiste à consigner les informations portant sur le processus et les résultats des discussions relatives à la hiérarchisation des besoins des différentes parties prenantes et aux décisions sur les questions d'évaluation.

Mettre en œuvre et adapter les critères au niveau institutionnel

Les définitions des critères et les présentes Orientations offrent un socle commun et un ensemble de définitions communément admises, à titre de point de départ. Il est néanmoins indispensable de les adapter au contexte institutionnel. Les évaluateurs et les responsables d'évaluation utilisent souvent les critères pour le compte d'une organisation, d'un ministère ou d'une autre institution en charge du développement, doté(e) de son propre mandat, de ses propres priorités stratégiques, de sa propre politique d'évaluation et de ses propres normes et orientations – autant d'éléments dont décident leurs organes de direction.

Lorsque les évaluateurs et les commanditaires étudient les moyens de mettre en œuvre et d'appliquer les critères, il importe qu'ils tiennent dûment compte, pour prendre leurs décisions, des priorités stratégiques et de la culture de l'organisation ainsi que des possibilités qu'elle offre. Certains termes (par exemple, « impact ») ne sont pas toujours employés de la même manière et il convient de prendre garde au risque de confusion ou de mauvaise interprétation des critères et de leur intention de départ. Une telle démarche aidera les évaluateurs et les commanditaires à appliquer les critères, permettra d'optimiser l'utilisation des constatations de l'évaluation et renforcera la pertinence par rapport aux besoins des utilisateurs cibles. Les responsables devraient encourager les échanges entre évaluateurs, commanditaires et public visé par l'évaluation de façon à examiner la manière dont les critères devraient être appliqués et interprétés. Un tel processus peut permettre de concevoir des évaluations crédibles et opportunes qui répondent mieux aux besoins des utilisateurs.

Les exigences d'une institution en ce qui concerne les méthodologies à suivre peuvent également avoir une incidence sur la manière dont les critères sont appliqués. Les évaluateurs devraient se référer aux exigences et orientations spécifiques de leur organisation/commanditaire. D'autres sources pertinentes telles que les orientations du Groupe des Nations Unies pour l'évaluation (GNUE) ou les normes de bonnes pratiques du Groupe de coopération pour l'évaluation (ECG) et les orientations du Réseau pour l'apprentissage actif sur la responsabilité et la performance au sein de l'action humanitaire (Active Learning Network for Accountability and Performance in Humanitarian Action, ALNAP) sur l'évaluation dans les contextes humanitaires sont également très utiles lorsqu'elles sont applicables.

Répondre aux exigences du Programme de développement durable à l'horizon 2030 et des Objectifs de développement durable

Outre la nécessité d'adapter les critères aux institutions qui les utilisent, la manière dont les critères sont compris et appliqués doit refléter le contexte général, ce qui influe sur leur utilisation par les responsables d'évaluation, les évaluateurs et les parties prenantes. Pour les dix prochaines années, ce sont les Objectifs de développement durable (ODD) et le Programme de développement durable à l'horizon 2030 qui constitueront le cadre d'action global et l'ensemble d'objectifs mondiaux les plus importants, en particulier pour les évaluateurs de la coopération internationale pour le développement.

Les principaux éléments du Programme 2030 sont notamment les suivants :

- l'accès universel aux fruits du développement ;
- l'inclusion, en particulier des personnes qui risquent le plus d'être laissées de côté ;
- les droits humains, l'égalité femmes-hommes et d'autres considérations d'équité ;
- la durabilité environnementale, le changement climatique et la gestion des ressources naturelles ;
- la complexité du contexte et des interventions à l'appui du développement ;
- les synergies entre les acteurs engagés dans le processus de développement.

Ce cadre influence à la fois les interventions et leur évaluation, y compris la manière dont les critères sont interprétés, ainsi que le processus d'évaluation proprement dit (notamment les personnes qui sont chargées d'appliquer les critères et de définir des questions prioritaires). L'Encadré 3.3 donne des orientations sur la façon dont le Programme 2030 peut être utilisé pour définir des programmes d'évaluation nationaux et l'Encadré 3.4 montre comment le système allemand d'évaluation du développement y est parvenu. Des efforts analogues ont été déployés par plusieurs États – dont le Costa Rica, le Nigéria et la Finlande – et par des organisations non gouvernementales (ONG), qui livrent des enseignements utiles pour les évaluateurs et les responsables d'évaluation (D'Errico, Geoghe et Piergallini, 2020^[9]).

Les définitions révisées des critères prennent également en compte ces éléments à plusieurs égards, par exemple en accordant une importance particulière au contexte et aux points de vue et priorités des bénéficiaires pour examiner la pertinence, l'efficacité et l'impact ; en tenant compte de l'équité des résultats pour analyser l'efficacité et l'impact ; et en adoptant une approche intégrée pour évaluer la cohérence. Les principes directeurs tiennent également compte du Programme 2030 en prônant un mode de réflexion intégré.

Encadré 3.3. S'appuyer sur les ODD pour définir des programmes d'évaluation nationaux

La publication intitulée « L'évaluation pour relier les priorités nationales aux ODD : Un guide pour les commanditaires et responsables d'évaluation » expose cinq aspects à prendre en compte pour aider les pays à concevoir des programmes d'évaluation et à renforcer l'utilité des évaluations. Les évaluateurs sont encouragés à appliquer la « théorie des systèmes complexes » et à :

- penser au-delà des politiques, programmes et projets individuels ;
- examiner les forces au niveau macro qui influent sur les chances de réussite ou d'échec ;
- avoir une compréhension nuancée de la notion de « réussite » ;
- reconnaître l'importance de la culture ;
- faire leur l'approche de réflexion évaluative et la gestion adaptative.

Source : Ofir et al. (2016^[9]), *Briefing: Five considerations for national evaluation agendas informed by the SDGs*, <https://pubs.iied.org/sites/default/files/pdfs/migrate/17374IIED.pdf>

Encadré 3.4. Relier les critères au Programme 2030 et aux ODD

Les questions qui suivent ont été élaborées par le ministère fédéral allemand de la Coopération économique et du développement (BMZ), avec le concours de l'Institut allemand pour l'évaluation du développement (DEval), afin d'aider les évaluateurs à déterminer quelle contribution globale a été apportée à la réalisation du Programme 2030 et des ODD. Chaque question correspond à des principes consacrés par les ODD et se rapporte également aux critères.

Universalité, responsabilité commune et redevabilité

- Dans quelle mesure l'intervention contribue-t-elle à la réalisation des ODD ? (voir critère d'impact)
- Dans quelle mesure l'intervention est-elle conçue de manière à utiliser les structures et les systèmes existants (des partenaires/d'autres donateurs/d'organisations internationales) pour mettre en œuvre les activités et dans quelle mesure ces structures et systèmes sont-ils utilisés ? (voir critère de cohérence)
- Les tâches sont-elles partagées avec d'autres donateurs et partenaires au développement pour mettre en œuvre l'intervention ? Si oui, dans quelle mesure ? (voir critère de cohérence)
- Dans quelle mesure les systèmes communs sont-ils utilisés pour le suivi, l'apprentissage et l'exercice de la redevabilité ? (voir critère de cohérence)

Interaction entre les dimensions économique, environnementale et sociale du développement

- Dans quelle mesure l'intervention adopte-t-elle une approche globale du développement durable (dimensions sociale, environnementale et économique) ? (voir critère de pertinence)
- Dans quelle mesure existait-il des interactions positives ou négatives, intentionnelles ou non, entre les réalisations sociales, économiques et environnementales, et quel a été l'impact global de l'intervention ? (voir critères d'efficacité et d'impact)
- Comment l'intervention a-t-elle contribué à favoriser les interactions positives ou négatives, intentionnelles ou non, entre les réalisations sociales, économiques et environnementales, et quel a été son impact global ? (voir critères d'efficacité et d'impact)

Inclusivité

- Dans quelle mesure l'intervention est-elle cohérente avec les normes et les critères internationaux relatifs à la participation et à la promotion des groupes particulièrement défavorisés et vulnérables ? (voir critère de cohérence)
- Dans quelle mesure y a-t-il eu des évolutions générales sur le plan du développement, positives ou négatives, intentionnelles ou non, au niveau des groupes particulièrement défavorisés et vulnérables (différenciation possible selon l'âge, le revenu, le sexe, l'origine ethnique, etc.) ? (voir critère d'impact)
- Comment l'intervention a-t-elle contribué à favoriser les évolutions générales sur le plan du développement, positives ou négatives, intentionnelles ou non, au niveau des groupes particulièrement défavorisés et vulnérables (différenciation possible selon l'âge, le revenu, le sexe, l'origine ethnique, etc.) ? (voir critère d'impact)
- Dans quelle mesure l'intervention a-t-elle contribué au renforcement de la résilience des groupes particulièrement défavorisés ou vulnérables (différenciation possible selon l'âge, le revenu, le sexe, l'origine ethnique, etc.) ? (voir critère de viabilité/durabilité)

Source : BMZ et DEval (2020^[10]), *Evaluation Criteria for German Bilateral Development Co-operation*

Appliquer un prisme sexospécifique aux critères

Les évaluateurs devraient adopter des méthodes de travail qui prennent en compte de manière réfléchie les différences d'expérience et d'impact selon le sexe, ainsi que leur interaction avec d'autres formes de discrimination dans un contexte donné (fondées, par exemple, sur l'âge, la « race » et l'origine ethnique, le statut social). Quelle que soit l'intervention, les évaluateurs devraient s'intéresser à la manière dont les dynamiques de pouvoir fondées sur le sexe se croisent et interagissent avec d'autres formes de discrimination d'une manière qui influe sur la mise en œuvre et les résultats de l'intervention. Il peut s'agir d'étudier la façon dont l'économie politique et le contexte socioculturel dans lequel s'inscrivent les interventions influencent la poursuite et la réalisation des objectifs.

Le fait d'appliquer un prisme sexospécifique peut livrer des données factuelles à l'appui de l'apprentissage et de l'exercice de la redevabilité, tout en favorisant la réalisation des objectifs liés à l'égalité des sexes. Pour ce faire, il convient notamment de prendre les mesures concrètes suivantes :

- adopter, pour les évaluateurs, les responsables et les commanditaires, des méthodes de travail inclusives qui favorisent une participation appropriée à la prise de décisions, à la collecte de données, à l'analyse et à la diffusion des constatations ;
- examiner la mesure dans laquelle, dans le cadre de l'intervention, le genre interagit avec d'autres barrières sociales pour compromettre l'égalité des chances ;
- examiner la manière dont une intervention interagit avec le contexte législatif, économique, politique, religieux et socioculturel pour mieux interpréter les différences en termes d'expérience des parties prenantes et en termes d'impact ;
- prendre en compte les constructions sociales de la masculinité et de la féminité ainsi que toute évolution de la dynamique et des rôles entre les sexes ;
- lors de la sélection des évaluateurs, analyser leurs compétences en matière d'approches d'évaluation tenant compte des sexospécificités ainsi que leur expérience acquise dans des contextes de travail diversifiés.

Le tableau ci-après vise à aider les évaluateurs à réfléchir sur la manière dont ils peuvent appliquer un prisme sexospécifique aux critères :

Tableau 3.1. Appliquer un prisme sexospécifique aux critères

| Critères | Questions pour guider les évaluateurs |
|----------------------|---|
| Pertinence | <p>L'intervention a-t-elle été conçue de manière à répondre aux besoins et aux priorités des différents sexes ? Si oui, comment ?</p> <p>Dans quelle mesure la conception de l'intervention tient-elle compte des droits des personnes de tout sexe et prévoit un retour d'information de diverses parties prenantes locales, y compris les groupes marginalisés ?</p> <p>L'intervention répond-elle aux besoins concrets et stratégiques des différents sexes ?</p> |
| Cohérence | <p>Dans quelle mesure la conception, la mise en œuvre et les résultats de l'intervention sont-ils cohérents avec la législation et les engagements internationaux en faveur de l'égalité et des droits femmes-hommes, notamment la Convention des Nations Unies sur l'élimination de toutes les formes de discrimination à l'égard des femmes (CEDAW), la Déclaration et le Programme d'action de Pékin, le Programme d'action de la Conférence internationale sur la population et le développement et le Programme de développement durable à l'horizon 2030 ?</p> <p>Dans quelle mesure l'intervention soutient-elle la législation et les initiatives nationales visant à améliorer l'égalité des sexes et les droits de la personne ? Quels enseignements peuvent être tirés ?</p> |
| Efficacité | <p>L'intervention a-t-elle atteint ses objectifs et résultats attendus d'une manière qui contribue à l'égalité des sexes ? Si oui, comment ?</p> <p>Les résultats ont-ils été différents selon le groupe de personnes ? Si oui, comment et pourquoi ? Des approches différentes ont-elles été nécessaires pour atteindre les personnes de différents sexes ? Le suivi et l'analyse des effets différenciés ont-ils été suffisants ? L'intervention a-t-elle été adaptée pour répondre aux éventuelles préoccupations et maximiser son efficacité ?</p> <p>La théorie du changement et le cadre de résultats reposaient-ils sur une analyse de la situation en matière d'égalité entre les sexes, sur une analyse de l'économie politique et sur les droits humains ? Si oui, dans quelle mesure ?</p> <p>Dans quelle mesure l'efficacité varie-t-elle selon le sexe des personnes et pourquoi ?</p> |
| Efficience | <p>Des ressources différentes ont-elles été allouées d'une manière tenant compte de l'égalité entre les sexes ? Si oui, de quelle manière ont-elles été allouées ? L'allocation différenciée des ressources était-elle appropriée ?</p> <p>Les coûts d'investissement par personne ciblée sont-ils en adéquation avec les besoins différenciés des personnes de sexes différents ?</p> |
| Impact | <p>Les impacts ont-ils été les mêmes pour les différents sexes ou y avait-il des différences en termes d'engagement, d'expérience et d'impact ? Si oui, comment s'expliquent ces différences d'impact ?</p> <p>Dans quelle mesure les impacts liés au sexe se croisaient-ils avec d'autres barrières sociales, notamment la « race »/l'origine ethnique, le handicap, l'âge et l'orientation sexuelle, de manière à contribuer à des différences en termes d'expérience et de réalisation ?</p> <p>Comment les normes et les barrières sexuées en vigueur dans le contexte politique, économique, religieux, législatif et socioculturel plus vaste ont-elles influé sur les réalisations ?</p> <p>Dans quelle mesure l'impact a-t-il contribué à des rapports de force égaux entre les personnes de sexe différent ainsi qu'à l'évolution des normes et des systèmes sociaux ?</p> |
| Viabilité/durabilité | <p>L'intervention a-t-elle contribué à l'égalité entre les sexes au sein des systèmes juridique, politique, économique et social plus vastes ? Si oui, de quelle manière et dans quelle mesure ? Cela s'est-il traduit par des changements durables au niveau de normes sociales préjudiciables pour les personnes de tous sexes ou d'un sexe donné ? Si non, pourquoi pas ?</p> <p>Les réalisations en matière d'égalité entre les sexes perdureront-elles une fois l'intervention achevée ? Les processus ont-ils contribué à pérenniser ces acquis ? Des mécanismes ont-ils été mis en place pour promouvoir la réalisation de l'égalité entre les sexes sur le long terme ?</p> |

Utilisation d'autres critères

Les six critères ont vocation à former un ensemble complet reflétant pleinement toutes les notions importantes qui devraient être couvertes par les évaluations. S'ils sont appliqués de manière réfléchie et

en tenant compte du contexte, ils conviendront pour des évaluations dans le domaine du développement durable comme de l'action humanitaire.

Pour autant, dans certains contextes, d'autres critères sont utilisés. Ainsi, dans leurs politiques d'évaluation, de nombreuses institutions commandent des analyses de domaines d'intérêt particuliers. En 2020, une évaluation des programmes menés par l'Italie en Bolivie dans le secteur de la santé a recouru à neuf critères, à savoir : la pertinence, l'efficacité, l'efficience, l'impact, la viabilité/durabilité, la cohérence, la valeur ajoutée de la coopération italienne, la visibilité de la coopération italienne et l'appropriation (Eurecna Spa, 2020^[11]). Dans un autre exemple, il s'agissait d'appliquer les critères dans des situations humanitaires, où des critères comme l'adéquation, la portée et l'existence de liens sont particulièrement pertinents³. L'additionnalité est un autre critère qui est parfois appliqué, souvent en relation avec le financement mixte, les financements non souverains et le financement climatique. Différentes définitions de ce terme, y compris différents types d'additionnalité financière et non financière, sont utilisés⁴. Selon la définition appliquée, l'additionnalité peut être évaluée au regard du critère de pertinence, d'efficacité ou d'impact. Elle est parfois elle-même considérée comme un critère transversal.

Les utilisateurs devraient faire preuve de prudence lorsqu'ils envisagent d'ajouter des critères, car une telle démarche peut créer la confusion et conduire à une évaluation trop générale (et donc à une analyse moins utile). S'appuyer sur un nombre limité de critères est utile pour permettre une analyse suffisamment approfondie et garantir la clarté conceptuelle – un point souligné à de nombreuses reprises au cours des consultations menées en vue de l'actualisation des définitions, en 2017-19.

Si l'on décide de recourir à d'autres critères, il importe alors de les définir. Justifier leur ajout peut permettre de comprendre comment ils se situent par rapport aux six critères décrits dans le présent document. Pour favoriser l'apprentissage d'une évaluation à l'autre, il est impératif d'évaluer les mêmes notions ou éléments au regard des mêmes critères.

Quels que soient les critères utilisés, les orientations et les principes fondamentaux décrits dans ce document devraient être appliqués.

Références

- ADE (2019), *Joint Evaluation of the Integrated Solutions Model in and around Kalobeyei, Turkana, Kenya*, HCR et DANIDA, https://um.dk/en/danida-en/results/eval/eval_reports/publicationdisplaypage/?publicationid=dd54bef1-1152-468c-974b-c171fbc2452d (consulté le 11 janvier 2021). [4]
- Bamberger, M., J. Vaessen et E. Raimondo (2015), *Dealing With Complexity in Development Evaluation - A Practical Approach*, https://www.betterevaluation.org/en/resources/dealing_with_complexity_in_development_evaluation (consulté le 11 janvier 2021). [13]
- BMZ et DEval (2020), *Evaluation Criteria for German Bilateral Development Co-operation*. [10]
- Bryld, E. (2019), *Evaluation of Sida's Support to Peacebuilding in Conflict and Post-Conflict Contexts: Somalia Country Report*, Asdi, https://publikationer.sida.se/contentassets/1396a7eb4f934e6b88e491e665cf57c1/eva2019_5_62214en.pdf (consulté le 11 janvier 2021). [5]
- D'Errico, S., T. Geoghe et I. Pierrgallini (2020), *Evaluation to connect national priorities with the SDGs*, IIED, <https://pubs.iied.org/17739IIED> (consulté le 22 février 2021). [8]

- Danida (2019), *Evaluation of Water, Sanitation and Environment Programmes in Uganda (1990-2017)*, Evaluation Department, Ministère des Affaires étrangères du Danemark, <http://www.oecd.org/derec/denmark/denmark-1990-2017-wash-environment-uganda.pdf> (consulté le 11 janvier 2021). [3]
- Davis, R. (2013), « Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations », n° 40, DIFD, <https://www.gov.uk/government/publications/planning-evaluability-assessments> (consulté le 12 janvier 2021). [12]
- Eurecna Spa (2020), *Bolivia - Evaluation of Health Initiatives (2009-2020)*, Ministère des Affaires étrangères et de la Coopération internationale de l'Italie, http://www.oecd.org/derec/italy/evaluation-report-of-health-initiatives-in-Bolivia-2009_2020.pdf (consulté le 11 janvier 2021). [11]
- Helle, E. et al. (2018), *From Donors to Partners? Evaluation of Norwegian Support to Strengthen Civil Society in Developing Countries through Norwegian Civil Society Organisations*, Agence norvégienne de coopération pour le développement (NORAD), https://www.norad.no/globalassets/filer-2017/evaluating/1.18-from-donor-to-partners/1.18-from-donors-to-partners_main-report.pdf (consulté le 11 janvier 2021). [2]
- Molund, S. et G. Schill (2004), *Looking Back, Moving Forward Sida Evaluation Manual*, Asdi, <https://www.oecd.org/derec/sweden/35141712.pdf> (consulté le 11 janvier 2021). [6]
- OCDE (2019), *Des meilleurs critères pour des meilleurs évaluations : Définitions adaptées et principes d'utilisation*, Réseau du CAD sur l'évaluation du développement, Éditions OCDE, Paris, <https://www.oecd.org/fr/cad/evaluation/criteres-adaptees-evaluation-dec-2019.pdf> (consulté le 11 janvier 2021). [1]
- OCDE (2002), *L'évaluation et l'efficacité de l'aide n° 6 - Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats (en anglais, français et espagnol)*, Éditions OCDE, Paris, <https://doi.org/10.1787/9789264034921-en-fr>. [14]
- Ofir, Z. et al. (2016), *Briefing: Five considerations for national evaluation agendas informed by the SDGs*, IIED, Londres, <https://doi.org/10.3138/cjpe.30.3.02.11>. [9]
- PAM (2016), *Technical Note: Evaluation Methodology*, DEQAS, Programme alimentaire mondial, <https://docs.wfp.org/api/documents/704ec01f137d43378a445c7e52dcf324/download/> (consulté le 11 janvier 2021). [7]

Notes

¹ Par « évaluabilité » est entendu la mesure dans laquelle une activité ou un programme est susceptible d'être évalué(e) d'une manière fiable et crédible. L'analyse de l'évaluabilité nécessite d'examiner à un stade précoce une activité proposée afin de déterminer si ses objectifs sont correctement définis et ses résultats vérifiables (OCDE, 2002^[14]). Voir également Davis (2013^[12]).

² Le processus d'élaboration des questions d'évaluation et de dialogue avec les parties prenantes, en tenant compte également de la complexité de l'intervention, est examiné plus avant, par exemple, dans Bamberger, Vaessen et Raimondo (2015^[13]).

³ Le réseau ALNAP a entrepris d'actualiser ses orientations de 2006 sur l'utilisation des critères dans les contextes humanitaires, en complément de son guide complet sur les évaluations relatives à l'action humanitaire (*Evaluation of Humanitarian Action Guide*).

⁴ Le groupe de travail d'EvalNet sur l'évaluation du financement mixte a contribué aux recherches sur les définitions de l'additionnalité et de notions connexes. Ses conclusions ont été publiées en 2021.

4 Comprendre les six critères : définitions, éléments d'analyse et principaux défis

Ce chapitre décrit plus précisément chacun des six critères. Il étudie ce que leur définition signifie en pratique, présente différents éléments d'analyse et explique comment chaque critère peut être appliqué d'une manière qui tienne décisivement compte des dimensions relatives à l'inclusion et à l'égalité. Pour chacun des critères, un tableau décrit les principales difficultés que pose leur application et formule des recommandations concrètes sur la façon dont les évaluateurs et les responsables d'évaluation peuvent surmonter ces difficultés. Le chapitre est ponctué de divers exemples qui montrent comment les critères ont été appliqués dans la pratique et favorisent la réflexion critique sur les dimensions de chaque définition.

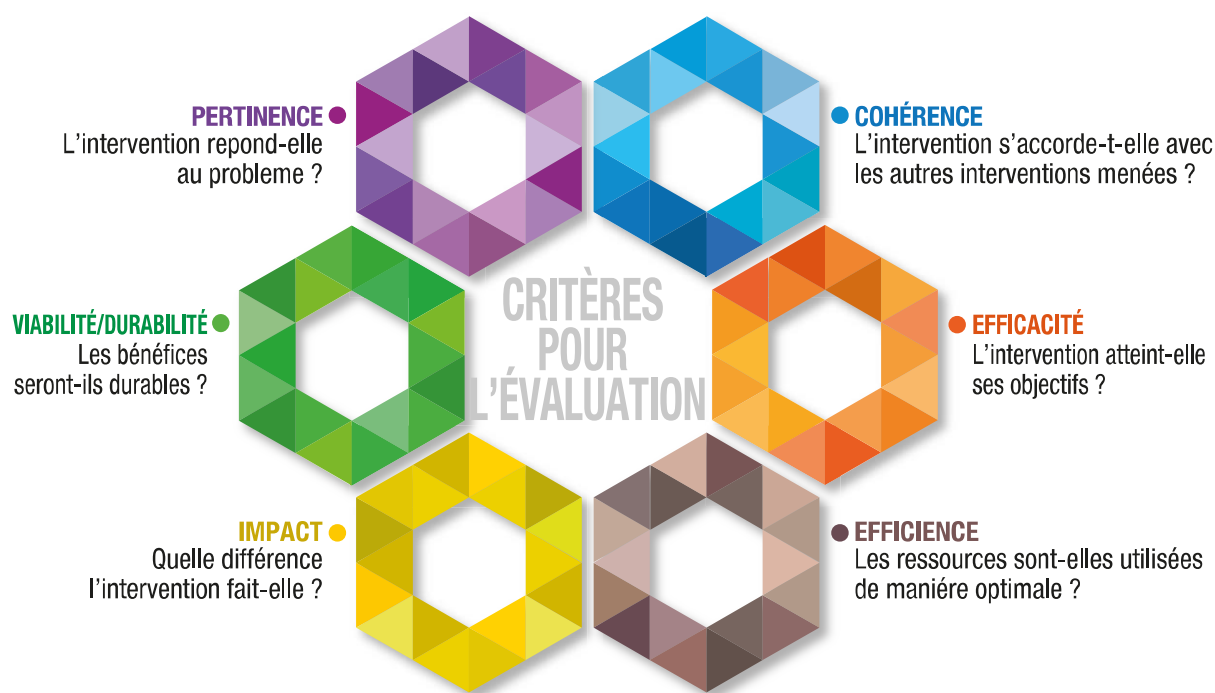
Ce chapitre présente les six critères (Graphique 4.1) dans l'ordre dans lequel il est le plus logique de les examiner : la pertinence et la cohérence d'abord, puis l'efficacité et l'efficience, et enfin l'impact et la viabilité/durabilité. Il définit chaque critère et en décrit l'importance. Il explique ensuite plus précisément chaque définition en examinant leurs différents éléments d'analyse, c'est-à-dire les principales notions qui les sous-tendent. Il ne s'agit pas de sous-critères, mais d'une illustration des diverses manières dont un critère peut être appliqué en fonction du contexte et de l'objectif visé. Les liens avec les autres critères sont également examinés.

Le chapitre contient aussi un début de réflexion sur des aspects essentiels liés à l'inclusion et au principe de « ne laisser personne de côté », consacré par le Programme de développement durable à l'horizon 2030. D'autres travaux sont en cours pour étudier les moyens d'appliquer un prisme sexospécifique aux critères et aux approches d'évaluation, notamment l'élaboration d'orientations spécifiques sur l'égalité femme-homme, l'autonomisation des femmes et les droits humains.

Pour chaque critère, un tableau décrit les difficultés couramment rencontrées et propose des solutions pour aider les évaluateurs et les responsables d'évaluation à les surmonter. Ces tableaux seront actualisés au fil du temps, à mesure que seront partagées les expériences acquises dans l'application des nouvelles définitions.

Enfin, le chapitre fournit des exemples réels qui montrent comment chaque critère peut être interprété.

Graphique 4.1. Les six critères et leurs questions associées



Pertinence

Définition de la pertinence :

Pertinence : l'intervention répond-elle au problème ?

Mesure dans laquelle les objectifs et la conception de l'intervention correspondent aux besoins, aux politiques et aux priorités des bénéficiaires¹, du pays, de la communauté internationale et des partenaires/institutions et demeurent pertinents même si le contexte évolue.

Remarque : l'expression « correspondent aux » signifie que les objectifs et la conception de l'intervention tiennent compte des conditions – économiques, environnementales, d'équité, sociales, liées à l'économie politique et aux capacités – dans lesquelles l'intervention est menée. Le terme « partenaires/institutions » englobe les administrations (nationales, régionales, locales), les organisations de la société civile, les entités privées et les organismes internationaux participant au financement, à la mise en œuvre et/ou à la supervision de l'intervention. L'évaluation de la pertinence consiste à examiner les différences et les arbitrages entre des priorités ou des besoins différents. Elle exige également d'analyser toute évolution du contexte pour pouvoir déterminer dans quelle mesure l'intervention peut être (ou a été) adaptée afin de conserver sa pertinence.

Qu'est-ce que la pertinence et pourquoi est-elle importante ?

L'évaluation de la pertinence aide les utilisateurs à comprendre si une intervention répond réellement au problème identifié. Elle permet aux évaluateurs de déterminer si les objectifs et la mise en œuvre d'une intervention correspondent vraiment aux besoins des bénéficiaires et des parties prenantes, de même qu'aux priorités sous-tendant l'intervention. Elle permet d'établir si les parties prenantes ciblées perçoivent l'intervention comme utile et importante.

La pertinence est une considération qui revêt un intérêt tout au long du programme ou du cycle de l'action publique, depuis la conception jusqu'à la mise en œuvre. Elle peut également être examinée sous l'angle d'objectifs mondiaux tels que les Objectifs de développement durable (ODD). Elle peut être appréhendée à travers quatre éléments d'analyse potentiels : la pertinence au regard des besoins des bénéficiaires et des parties prenantes, la pertinence au regard du contexte, la pertinence de la qualité et de la conception et la pertinence au fil du temps. Ces éléments sont examinés plus en détail dans la partie qui leur est consacrée. Ils devraient être intégrés à l'analyse en fonction de l'objectif de l'évaluation, mais ne sont pas exhaustifs.

Pour évaluer la pertinence, il faut commencer par déterminer si les objectifs de l'intervention ont été correctement définis, s'ils sont réalistes et réalisables, et si les résultats sont vérifiables et en phase avec les normes internationales applicables aux interventions de développement – une démarche qui devrait être compatible avec la notion d'évaluabilité décrite dans les *Normes de qualité pour l'évaluation du développement de l'OCDE* (OCDE, 2010^[1]). Les résultats ou l'énoncé des objectifs peuvent être formulés de manière inintelligible ou trop vague, être difficiles à mesurer ou être axés sur les activités ou les intrants. Dans certains cas, il est nécessaire de redéfinir ou de reconstruire la théorie du changement pour que l'évaluateur identifie clairement ces objectifs. Les évaluateurs devraient veiller à effectuer leur évaluation au regard d'objectifs réalistes et de bonne qualité. Par ailleurs, les indicateurs utilisés pour mesurer la réalisation des objectifs devraient être validés sur la base de critères généralement admis comme les indicateurs SMART (spécifique, mesurable, atteignable, réaliste et défini dans le temps) (IDD and Associates, 2006^[2]). Les évaluations qui prennent en compte le critère de la pertinence devraient s'attacher à déterminer si les objectifs ont été mis en œuvre d'une manière conforme aux bonnes pratiques, et

comment. Elles devraient aussi analyser les capacités organisationnelles et les compétences des partenaires d'exécution, ainsi que la manière dont ils se sont adaptés aux éventuelles évolutions du contexte.

Les évaluateurs devraient également identifier clairement les parties prenantes dont les besoins et les priorités devraient être pris en compte pour l'évaluation de la pertinence – les bénéficiaires, mais aussi les partenaires et les institutions qui participent au financement, à la supervision ou à la mise en œuvre de l'intervention. Il faudrait accorder une importance particulière aux bénéficiaires. En effet, l'appropriation d'une intervention est importante, et les bénéficiaires sont considérés comme les principales parties prenantes dans la définition des priorités et des besoins. En fonction de l'intervention, il peut aussi être justifié de tenir compte des besoins nationaux et infranationaux (le cas échéant), des stratégies locales mises en œuvre pour répondre aux besoins et du degré de concordance entre l'intervention et ces besoins. Les besoins institutionnels peuvent englober les besoins des donateurs, mais ne s'y limitent pas. En conséquence, il est possible d'évaluer la pertinence dans divers contextes, y compris lorsqu'il n'existe pas de donneur proprement dit, mais un ensemble de partenaires d'intervention (par ex. dans le cas d'une politique commerciale).

Comprendre la pertinence : éléments d'analyse

La définition de la pertinence recouvre quatre grandes dimensions : la prise en compte des besoins, des politiques et des priorités ; la sensibilité et l'adaptation au contexte ; la qualité de la conception ; et la capacité d'adaptation au fil du temps.

Prise en compte des besoins, des politiques et des priorités

Pour analyser la pertinence d'une intervention, l'élément le plus important, peut-être, consiste à déterminer dans quelle mesure elle répond aux besoins et aux priorités des bénéficiaires. Une telle analyse permet de comprendre à quelles problématiques l'intervention vise à répondre, et pourquoi. Les bénéficiaires sont des parties prenantes essentielles et devraient être pris en compte tout au long de l'intervention. Il ne s'agit pas nécessairement des destinataires directs des services. Selon le type d'intervention, ils peuvent se situer (beaucoup) plus en amont dans la chaîne de résultats. Ainsi, une intervention peut avoir pour but d'accroître les capacités d'un bureau national d'audit. Ce renforcement des capacités améliorera la gestion des finances publiques et, partant, contribuera à la réalisation d'ODD, par exemple dans le domaine de la santé et de l'éducation. Toutefois, dans le cadre de l'évaluation de ce soutien au renforcement des capacités, ce sont les membres du personnel du bureau d'audit qui seront considérés comme les principaux bénéficiaires. Désigner clairement les bénéficiaires (principaux et secondaires) est une première étape indispensable de l'évaluation de la pertinence.

Analyser les besoins des bénéficiaires et examiner s'ils sont pris en compte renseigne non seulement sur l'aptitude à répondre à ces besoins, mais aussi sur le niveau d'appropriation et de participation concernant la conception et la mise en œuvre de l'intervention (ce qui peut influencer sur d'autres critères). Cela permet de comprendre qui est associé à la conception d'une intervention et qui en est exclu, et, partant, quelle incidence cette situation a sur la conception et la mise en œuvre de l'intervention.

Le critère de pertinence suppose de se concentrer sur les domaines où les besoins sont les plus grands ou, pour reprendre les termes du Programme 2030, sur l'objectif d'aider en premier lieu les plus défavorisés. De fait, il est particulièrement utile pour comprendre qui participe à une intervention et qui est touché par une intervention. Il offre aux évaluateurs la possibilité d'examiner si et à quel point les publics marginalisés sont pris en compte dans les priorités stratégiques et dans celles de l'intervention. Même une intervention parfaitement en phase avec la politique en vigueur peut être en décalage avec les priorités concrètes des participants, qui peuvent ne pas avoir été associés à la définition des priorités et plans officiels.

Pour évaluer la pertinence, il faut aussi tenir compte de la manière dont l'intervention répond aux priorités des institutions ou partenaires engagés – administrations (nationales, régionales, locales), organisations de la société civile, entités privées et organismes internationaux participant au financement, à la mise en œuvre et/ou à la supervision de l'intervention. Il s'agit d'examiner le degré d'alignement sur les stratégies et les politiques de ces institutions.

Pour évaluer la pertinence d'une intervention par rapport aux politiques, aux priorités et aux besoins mondiaux, il convient d'examiner sa contribution à la réalisation d'objectifs généraux de portée mondiale (c'est-à-dire l'importance relative de l'intervention par rapport à l'ensemble des efforts déployés), ce qui suppose souvent, pour les évaluateurs, de comparer les résultats (potentiels) obtenus dans le contexte/pays donné à d'autres résultats obtenus ailleurs. Ces questions relatives à la pertinence d'un point de vue mondial ne sont pas toujours examinées au stade de la conception de l'intervention. C'est pourquoi les évaluateurs peuvent apporter des analyses utiles en la matière, qui peuvent aider les parties prenantes à appréhender l'importance stratégique d'une intervention au-delà de son contexte local.

La définition de la pertinence exige aussi des évaluateurs qu'ils examinent les éventuelles tensions ou arbitrages entre les différents publics dont les priorités et les besoins sont visés par l'intervention. Les points de vue des participants et d'autres parties prenantes peuvent diverger et l'évaluation devra donc mettre à plat ces différences et examiner les conséquences des choix effectués. À titre d'exemple, l'évaluation d'une intervention ayant pour but d'éliminer une maladie – par exemple la poliomyélite – de la surface du globe examinera la charge de morbidité relative sous l'angle mondial pour déterminer dans quel pays ou région il est prioritaire d'agir. Dans certains cas, cette priorité mondiale (par exemple, vacciner contre la poliomyélite dans la dernière région où la maladie circule au sein de la population) peut être en contradiction avec les priorités locales (les bénéficiaires accordant, par exemple, la priorité aux problèmes d'approvisionnement en eau et d'assainissement). Il peut être utile pour les évaluateurs de mettre au jour ces tensions à travers une analyse rigoureuse de la pertinence.

Sensibilité et adaptation au contexte

Les besoins des bénéficiaires et d'autres parties prenantes importantes ne peuvent pas être appréhendés en dehors du contexte dans lequel ils s'inscrivent et par lequel ils sont déterminés. Ce contexte est multidimensionnel et fonction des facteurs suivants : économique, environnemental, lié à l'équité, social, culturel, lié à l'économie politique et considérations relatives aux capacités. Les évaluateurs sont encouragés à déterminer quels facteurs contextuels sont les plus pertinents pour une intervention donnée.

La pertinence du contexte peut être analysée aussi bien au niveau de la conception que de la mise en œuvre de l'intervention. La prise en compte du contexte diffère également selon que l'évaluation se fait *ex ante* ou *ex post*. Les évaluateurs peuvent par exemple se poser des questions au sujet de la manière dont le contexte a été compris et pris en compte au stade de la conception. S'agissant des évaluations *ex post*, les évaluateurs devraient se poser la question de savoir si le contexte a changé entre le début et la fin de l'intervention. Les évaluations *ex post* disposent de davantage d'éléments contextuels sur lesquels se fonder et devraient en tenir compte dans leur analyse. La dimension contextuelle est complémentaire de la dimension temporelle dans l'analyse de la pertinence d'une intervention, puisqu'elle prend en compte les éventuelles variations de la pertinence à mesure que la situation évolue.

Le contexte historique peut également être pris en considération. Il s'agit par exemple de se demander si des interventions similaires ont déjà été mises en œuvre par le passé ou s'il existe des tensions historiques ou des antécédents législatifs ou politiques susceptibles d'avoir une incidence sur la compréhension des besoins et la définition des objectifs. Le contexte historique peut également englober les hypothèses faites par le passé au sujet de la pertinence d'une intervention. Il s'agit alors de vérifier si ces hypothèses sont toujours valables. Lorsque des évaluations ont déjà été conduites, ces hypothèses peuvent aider à retracer le contexte historique et à déterminer si les interventions tirent parti des leçons des évaluations précédentes.

Qualité de la conception

L'évaluation de la « qualité de la conception » permet d'examiner dans quelle mesure l'intervention a été conçue pour répondre à des priorités et des besoins pertinents et si les objectifs ont été définis avec précision. Cette dimension permet en outre d'évaluer si les priorités et les besoins des parties prenantes sont clairement énoncés dans les objectifs de l'intervention, dans la théorie du changement sur laquelle elle se fonde, ainsi que dans sa théorie d'action et/ou son mode opératoire. Les évaluateurs sont ainsi en mesure de comprendre quelles lacunes, dans la conception du programme, ont pu réduire la pertinence globale d'une intervention. Cet élément d'analyse influe aussi sur l'évaluabilité de l'intervention en général parce qu'il implique que l'on s'intéresse également d'emblée à la qualité de la conception de l'intervention. Il apporte en outre un éclairage sur l'adéquation de l'intervention par rapport à l'institution qui la met en œuvre. Les évaluateurs peuvent par exemple se pencher sur la question de savoir si l'intervention a été conçue en tenant compte des questions de faisabilité technique, organisationnelle et financière.

Capacité d'adaptation au fil du temps

Les évaluateurs devraient examiner la manière dont les interventions peuvent évoluer au fil du temps. L'éclatement d'un conflit ou l'évolution du contexte politique et économique, par exemple, ont une forte incidence sur la mise en œuvre. L'évaluation de la pertinence au regard de la temporalité devrait comporter une analyse de la gestion adaptative. Les évaluateurs ne devraient pas se contenter d'analyser la pertinence au début et à la fin d'un programme, mais aussi examiner la manière dont ce dernier s'est adapté aux évolutions qui ont eu lieu tout au long de son cycle de vie. Une telle démarche permet d'analyser les éventuelles variations de la pertinence d'une intervention à mesure que les conditions de sa mise en œuvre évoluent. À titre d'exemple, il peut s'agir d'analyser dans quelle mesure les adaptations ont permis de continuer de répondre aux priorités et aux besoins les plus importants et si elles ont eu une incidence sur la qualité de la conception au fil du temps.

Là encore, l'adaptation continue au contexte extérieur et aux évolutions internes (par ex. lorsqu'une modification du financement nécessite un changement dans la programmation) devrait être prise en compte, de même que peuvent être considérés les risques et les opportunités, notamment la mesure dans laquelle le programme a atténué les risques pesant sur la réalisation de son objectif ou a été adapté pour saisir une chance ou mieux répondre aux besoins. L'adaptation peut amener à opérer des arbitrages entre les groupes dont les besoins sont jugés prioritaires, ce qui soulève des questions de redevabilité – un aspect qui devrait être pleinement étudié pour déterminer la manière dont la pertinence d'un programme a pu ou non s'en trouver modifiée.

Liens avec les autres critères

Les critères étant liés les uns aux autres, la pertinence peut avoir un lien avec les autres critères mobilisés pour l'évaluation. Elle est souvent considérée comme une condition sans laquelle les autres critères ne peuvent pas être satisfaits.

- L'évaluation de la pertinence fournit une base pour comprendre si les besoins sont satisfaits au regard des critères d'efficacité et d'impact. En effet, la pertinence en tant que critère est un préalable à l'efficacité, les besoins et les objectifs devant être clairement énoncés pour permettre d'évaluer cette dernière.
- La pertinence est complémentaire de la cohérence. Ces deux critères nécessitent une analyse contextuelle : dans le cas de la pertinence, pour appréhender l'adéquation entre l'intervention et les priorités et les besoins des principales parties prenantes ; et dans le cas de la cohérence, pour comprendre les liens avec d'autres interventions. La pertinence met l'accent sur la manière dont une intervention s'adapte au contexte. La cohérence élargit le prisme et examine d'autres interventions mises en œuvre dans le même contexte ainsi que leur interaction avec l'intervention

évaluée. Ensemble, la pertinence et la cohérence peuvent fournir une vision plus claire de l'influence que l'intervention a sur le contexte dans lequel elle est mise en œuvre – et vice-versa.

- L'analyse de la pertinence a aussi un lien avec l'impact, qui a trait à l'importance qu'une intervention a *in fine* – y compris l'intérêt qu'elle revêt pour les personnes touchées. Les évaluateurs devraient consacrer suffisamment de temps à l'examen des besoins, des priorités et des politiques des différents acteurs (y compris les tensions potentielles existantes) pour pouvoir évaluer de manière suffisamment approfondie la pertinence globale de l'intervention et analyser plus avant son importance lors de l'évaluation de l'impact.
- Bon nombre d'aspects de la pertinence sont également déterminants pour l'efficacité et la viabilité/durabilité. Ainsi, une intervention pertinente a davantage de chances de susciter l'adhésion des parties prenantes, ce qui peut avoir une incidence sur le calendrier d'exécution et l'utilisation des ressources, ainsi que sur le degré d'appropriation des bénéficiaires de l'intervention (et, partant, leur pérennité).

Intégrer la dimension de l'inclusion

Pour appréhender la pertinence d'une intervention, il est essentiel de comprendre la nature sexuée des dynamiques de pouvoir et de réfléchir sur l'engagement pris au titre des ODD de ne « laisser personne de côté ». Les dynamiques de pouvoir fondées sur le sexe et la marginalisation de certains groupes – notamment raciaux/ethniques – sont des considérations essentielles pour appréhender la pertinence dans un contexte donné.

Comprendre qui a été associé à la conception de l'intervention et de quelle manière, en attachant une attention particulière aux dynamiques de pouvoir et aux groupes marginalisés, permet aux évaluateurs de mieux appréhender la pertinence de l'intervention telle qu'elle a été conçue ainsi que la mesure dans laquelle elle s'est adaptée à l'évolution des besoins au fil du temps.

La définition de la pertinence présentée plus haut souligne combien il est important de considérer les arbitrages entre différents besoins et priorités et, notamment, de tenir davantage compte de l'équité et des dynamiques de pouvoir entre les personnes touchées directement ou indirectement par l'intervention.

De ce point de vue, il existe un lien fort avec les droits humains et l'égalité, en particulier si l'on adopte une approche intersectionnelle, c'est-à-dire si l'on observe comment diverses formes d'identités sociales et politiques – par exemple le genre, le handicap, l'origine ethnique, l'orientation sexuelle et la classe sociale – se conjuguent pour créer discriminations et inégalités. Lorsqu'on définit, dans le cadre de l'analyse de la pertinence, les priorités à poursuivre, il est essentiel de tenir compte des groupes sous-représentés et marginalisés (qui peuvent rencontrer des difficultés d'accès aux services et/ou aux droits) ainsi que de la manière dont leurs besoins et leurs priorités sont – ou non – pris en compte dans les politiques et les documents officiels. En outre, il importe d'examiner si l'intervention intègre différents niveaux d'accès, compte tenu des obstacles auxquels se heurtent certains groupes.

Difficultés posées par l'évaluation de la pertinence et solutions proposées

Le tableau ci-après décrit plusieurs difficultés couramment rencontrées lors de l'évaluation de la pertinence – éventail des besoins et des priorités à prendre en compte ; mauvaise formulation des objectifs ; évolution du contexte – et propose à l'intention des évaluateurs comme des responsables d'évaluation des solutions pour les surmonter.

Tableau 4.1. Difficultés posées par l'évaluation de la pertinence

| Difficulté | Solutions : évaluateurs | Solutions : responsables d'évaluation |
|--|---|--|
| Les parties prenantes nationales et internationales à prendre en compte sont nombreuses et ont de multiples priorités et besoins qui entrent potentiellement en concurrence. | Une analyse robuste de la pertinence est une analyse qui examine les arbitrages potentiels et effectifs à opérer pour répondre aux besoins. La pertinence des bénéficiaires devrait être le critère le plus important lors de l'examen des besoins et des priorités. L'appropriation et la participation permettent également de surmonter cette difficulté et de s'assurer que les priorités concurrentes sont prises en compte et, si possible, conciliées. | Définir de façon claire et succincte les bénéficiaires de l'intervention. Si la chaîne de résultats est très longue, mettre l'accent sur les bénéficiaires directs lors de l'analyse de la pertinence et réserver les questions relatives aux bénéficiaires finaux pour l'analyse de l'efficacité et de l'impact. Au moment de restreindre le champ d'analyse aux besoins et aux priorités les plus importants, tenir compte de la finalité de l'évaluation et déterminer quelle dimension de la pertinence peut être la plus utile pour y contribuer. Une attention particulière devrait être accordée à l'utilité de l'évaluation au moment de la conception. |
| Comprendre si une intervention poursuit des objectifs clairs et appropriés et comment ils ont été définis. | Les objectifs de l'intervention constituent un point de départ essentiel pour mener une évaluation robuste de la pertinence. Cette évaluation peut être difficile si ces objectifs ne sont pas clairement formulés. Les évaluateurs devraient procéder à un examen critique des objectifs pour déterminer s'ils sont suffisamment ambitieux et reflètent les besoins des bénéficiaires visés ainsi que les capacités des parties prenantes, et devraient étudier la manière dont ces objectifs ont été définis et vérifiés. L'adéquation des objectifs peut être analysée plus avant dans le cadre de l'évaluation de l'impact, qui permet de comprendre si les objectifs ont contribué à des changements importants et porteurs de transformations. Dans ce cas, les évaluateurs devraient établir un dialogue étroit avec les parties prenantes du programme de manière à comprendre quels sont les objectifs du programme et comment ils ont été définis, et à déterminer si un plan d'urgence est ou était en place. Ils devraient évaluer si et dans quelle mesure ces objectifs reposent sur une analyse solide des besoins et du contexte. Cette évaluation constituera un point de départ précieux pour appréhender la pertinence. | Les responsables d'évaluation devraient tenter d'établir les objectifs de l'intervention au moment de l'élaboration des termes de référence. Les termes de référence devraient permettre aux évaluateurs d'évaluer d'un œil critique le degré d'adéquation des objectifs par rapport aux besoins et au contexte extérieur ainsi que la prise en compte dans ces objectifs des commentaires des parties prenantes. |
| L'intervention n'est pas sous-tendue par une théorie du changement, un modèle logique et/ou une autre raison d'être explicite clairement énoncé(e). | L'évaluation de la pertinence est facilitée lorsqu'il existe une théorie du changement, un modèle logique et/ou une autre raison d'être explicite, qui apportent des éclairages importants pour évaluer la pertinence de la conception de l'intervention. Les évaluateurs devraient déterminer s'il existe de tels cadres et, le cas échéant, ils devraient coopérer avec les parties prenantes pour parvenir à reconstruire ou à énoncer clairement la théorie du changement appliquée, comme point de départ de l'évaluation. | Les responsables d'évaluation devraient s'efforcer de fournir aux évaluateurs les documents relatifs à la théorie du changement, au modèle logique et/ou à toute autre raison d'être explicite du programme/de l'intervention, dès le début de l'évaluation. |
| Le contexte a évolué au fil du temps, mais on a du mal à savoir si – et comment – l'intervention a été adaptée. | Malheureusement, les adaptations réalisées ne sont pas toujours dûment consignées. Il faut commencer par décrire les changements qui ont eu lieu, puis expliquer comment l'intervention a été adaptée (ou pas) pour pouvoir ensuite évaluer la pertinence. | Les responsables d'évaluation peuvent aider à identifier et à contacter les principaux décideurs, qui pourront apporter leur soutien pour « reconstituer » les changements apportés à l'intervention. En l'absence de décisions dûment consignées, |

| Difficulté | Solutions : évaluateurs | Solutions : responsables d'évaluation |
|------------|--|--|
| | S'il est évident que le contexte a connu des évolutions importantes susceptibles d'avoir influé sur la pertinence, les évaluateurs devraient consulter les documents relatifs à l'intervention (par exemple, rapports annuels ou trimestriels, données de suivi et, si possible, documents tels que procès-verbaux de réunion et communiqués au personnel et aux participants) de manière à « reconstituer » les changements qui ont été apportés et à identifier les principales décisions prises ainsi que leur justification. | des entretiens peuvent permettre de comprendre ce qui s'est passé. |

Exemples d'évaluation de la pertinence

Cette partie présente une sélection d'exemples d'évaluation de la pertinence tirés d'évaluations de programmes de soutien budgétaire général, de mesures d'aide aux moyens de subsistance des agriculteurs et de programmes dans le secteur de la santé.

Encadré 4.1. Évaluation de la pertinence des programmes de soutien budgétaire général

L'Approche méthodologique pour les évaluations de l'aide budgétaire élaborée par le Réseau du CAD de l'OCDE sur l'évaluation du développement (EvalNet) recommande de se concentrer sur la pertinence et l'adéquation de la conception du soutien budgétaire et des différents moyens mis en œuvre par rapport au contexte politique, économique et social du pays partenaire, au cadre d'action des pouvoirs publics et aux stratégies d'aide au développement des partenaires au développement.

Par exemple, une question essentielle qui peut être posée à cet égard est la suivante : dans quelle mesure la conception du programme de soutien budgétaire était-elle adaptée et pertinente compte tenu du contexte politique, économique et social du pays, du cadre d'action des pouvoirs publics et des stratégies d'aide au développement des partenaires extérieurs au développement ?

À titre d'illustration, dans le cadre de l'évaluation du soutien budgétaire en Tunisie (citée dans l'Approche méthodologique), les évaluateurs ont examiné dans quelle mesure le programme de soutien budgétaire avait recouru à la combinaison adéquate d'assistance technique, de dialogue sur les politiques et de financements compte tenu du contexte institutionnel de l'administration tunisienne et des priorités de celle-ci. En premier lieu, ils se sont penchés sur les réalisations au niveau de l'administration (en tant que bénéficiaire). Les réalisations au niveau des bénéficiaires finaux – les utilisateurs des services (par exemple les jeunes fréquentant l'enseignement supérieur) – ont été incluses dans une deuxième phase, consacrée à l'évaluation de l'efficacité et de l'impact.

Source : OCDE (2011^[3]), Approche méthodologique pour les évaluations de l'aide budgétaire, https://www.oecd.org/dac/evaluation/Approche%20methodologique_aide%20budgetaireFINAL.pdf

Encadré 4.2. Évaluation du projet d'amélioration des moyens de subsistance des agriculteurs dans la « zone sèche » du Myanmar

Cette intervention et son évaluation portaient sur les moyens de subsistance et la sécurité alimentaire des populations pauvres et vulnérables de la « zone sèche » du centre du Myanmar. L'évaluation fournit une bonne illustration de la manière dont le critère de pertinence peut être appliqué de façon adaptée au contexte et appropriée.

Il s'agissait notamment d'examiner dans quelle mesure l'intervention prenait en compte les besoins et les priorités des parties prenantes et comment les différents groupes de parties prenantes avaient été associés tout au long de sa conception et de sa mise en œuvre.

L'évaluation s'intéressait particulièrement au maintien de la pertinence au fil du temps et notamment à la capacité d'adaptation du programme face à des conditions évolutives pendant la mise en œuvre. Outre le maintien de la pertinence sur toute la durée du projet, l'évaluation s'est également arrêtée sur certains moments décisifs, comparant l'appréciation de la pertinence telle qu'elle ressortait de l'examen à mi-parcours avec celle de la pertinence au moment de l'évaluation.

Les évaluateurs ont conclu que l'intervention était très pertinente au regard des besoins et du contexte et qu'elle avait été particulièrement efficace dans son adaptation aux évolutions survenues au fil du temps dans les domaines de l'environnement et de l'action publique. L'évaluation a appliqué de manière adéquate le critère de pertinence en mobilisant et en examinant les besoins et le contexte à la fois au stade de la conception et pendant la mise en œuvre, ce qui permet d'avoir une compréhension complète de l'intervention et de son évolution.

Source : FAO (2020^[4]), *Evaluation of "Improving farmer livelihoods in the dry zone through improved livestock health, productivity and marketing"*, <http://www.fao.org/3/ca8463en/ca8463en.pdf>

Encadré 4.3. Évaluation de la pertinence des initiatives en faveur de la santé en Bolivie

Cette évaluation portait sur un programme de soutien au secteur de la santé en Bolivie mené sur la période 2009-20. Au titre de la pertinence étaient évaluées non seulement les demandes officielles approuvées du gouvernement bolivien, en particulier du ministère de la Santé, mais aussi la capacité d'adaptation à des besoins plus spécifiques des parties prenantes, dont les bénéficiaires. L'évaluation a dressé les conclusions suivantes :

Pertinence : Les projets correspondent aux initiatives et aux demandes des institutions et des autorités agréées par le ministère de la Santé, ce qui signifie que la coopération italienne apporte des réponses efficaces à ce que le gouvernement bolivien qualifie de nécessaire. Toutefois, il n'existe aucun processus approprié permettant d'identifier les besoins des projets dans leur ensemble, ce qui se traduit par des problèmes de pertinence et de cohérence. Les projets qui ont le mieux réussi parmi ceux évalués sont ceux pour lesquels un niveau au moins adéquat de connaissance des besoins spécifiques à satisfaire peut être démontré.

Source : Eurecna Spa (2020^[5]), *Bolivia - Evaluation of health initiatives (2009-2020)*, http://www.oecd.org/derec/italy/evaluation-report-of-health-initiatives-in-Bolivia-2009_2020.pdf

Cohérence

Définition de la cohérence :

Cohérence : l'intervention s'accorde-t-elle avec les autres interventions menées ?

Mesure dans laquelle l'intervention est compatible avec les autres interventions menées au sein d'un pays, d'un secteur ou d'une institution.

Remarque : le critère cherche à examiner comment d'autres interventions (en particulier des politiques) appuient ou affaiblissent l'intervention évaluée, et inversement. Sont englobées la cohérence interne et la cohérence externe. La cohérence interne concerne les synergies et les interdépendances entre l'intervention et les autres interventions menées par la même institution/administration, ainsi que la cohérence entre l'intervention et les normes et critères internationaux pertinents auxquels l'institution/administration adhère. La cohérence externe concerne la cohérence entre l'intervention considérée et celles menées par d'autres acteurs dans le même contexte. Elle englobe la complémentarité, l'harmonisation et la coordination avec les autres acteurs, et vérifie que l'intervention apporte une valeur ajoutée tout en évitant le chevauchement d'activités.

Qu'est-ce que la cohérence et pourquoi est-elle importante ?

Dans le contexte actuel, il convient d'accorder une attention accrue à la cohérence, en mettant davantage l'accent sur l'existence (ou l'absence) de synergies (ou d'arbitrages) entre les domaines de l'action publique, tout en prêtant plus d'attention à la coordination entre les différents secteurs de l'administration, en particulier lorsqu'il s'agit d'interventions humanitaires et de gestion de conflit, ou de faire face à l'urgence climatique.

Dans le droit fil du Programme 2030 et des ODD, ce nouveau critère encourage l'adoption d'une approche intégrée et offre un angle d'analyse important pour évaluer la cohérence en tenant compte des synergies, de la coordination entre administrations et de la conformité aux normes et critères internationaux. Il permet d'examiner divers arbitrages et tensions et de repérer les situations dans lesquelles les chevauchements d'activités ou les incohérences dans la mise en œuvre des politiques à l'échelle de l'administration ou de différentes institutions peuvent compromettre les avancées générales.

Ce critère encourage également les évaluateurs à comprendre quel rôle joue une intervention au sein d'un système particulier (organisation, secteur, domaine thématique, pays) au lieu de se concentrer uniquement sur l'intervention ou l'institution. Si l'analyse de la cohérence externe vise à comprendre dans quelle mesure les objectifs stratégiques des acteurs sont en adéquation avec les objectifs de développement internationaux, elle serait incomplète s'il était fait abstraction des intérêts, de l'influence et du pouvoir d'autres acteurs extérieurs. D'où l'utilité de se placer sous l'angle plus large de l'économie politique pour mieux appréhender la cohérence des interventions.

Par ailleurs, les sources de financement (internationales comme nationales) du développement durable sont de plus en plus diversifiées. La référence aux « normes et critères internationaux » dans la définition encourage l'analyse de la compatibilité de l'intervention avec les engagements souscrits par l'acteur en vertu du droit ou d'accords internationaux (lois contre la corruption, conventions sur les droits fondamentaux, etc.). Il s'agit des accords auxquels l'entité est déjà partie, donc d'une dimension qui entre dans le champ de la cohérence interne – aspect qui, jusqu'ici, faisait rarement l'objet d'une analyse suffisante. Le critère de cohérence peut également permettre d'évaluer les normes et les critères internationaux du point de vue de la prise en compte des priorités mondiales, ce qui constitue un angle d'analyse complémentaire.

Comprendre la cohérence : éléments d'analyse

La cohérence englobe la cohérence interne et la cohérence externe.

Cohérence interne

La cohérence interne des politiques recouvre deux éléments : la cohérence avec les cadres d'action plus larges des institutions et la cohérence avec d'autres interventions mises en œuvre par l'institution, y compris celles menées par d'autres services chargés de la mise en œuvre d'interventions de développement ou d'interventions susceptibles d'avoir une incidence sur le même contexte opérationnel². Il s'agit d'examiner dans quelle mesure ces activités sont harmonisées, si les efforts sont redondants ou s'il existe des chevauchements d'activités, et si les interventions sont complémentaires.

Au sein des administrations nationales (ou, le cas échéant, d'autres niveaux de gouvernement), il peut exister des problèmes de cohérence entre différents types de politiques publiques, entre différents niveaux de gouvernement et entre différentes parties prenantes (acteurs publics et non publics, commerciaux et non commerciaux). Ces aspects devraient être dûment pris en compte lors de l'évaluation de la cohérence afin de déterminer la place de l'intervention dans cet ensemble ainsi que son degré de compatibilité avec les politiques régissant le contexte en général.

À titre d'illustration, les lignes directrices pour l'évaluation de l'aide publique au développement (APD) élaborées par le ministère japonais des affaires étrangères (Ministère des affaires étrangères du Japon, 2019^[6]) aident les commanditaires d'évaluations et ceux qui sont chargés de mettre en œuvre ces évaluations à évaluer la cohérence des stratégies et des actions diplomatiques et en matière de développement menées par les différents secteurs de l'administration japonaise. Elles fournissent un cadre et des orientations qui aident les évaluateurs à prendre en compte les interconnexions, les relations de complémentarité et la cohérence entre les stratégies diplomatiques et en matière d'APD. Elles permettent ainsi d'analyser selon une approche globale l'engagement et le soutien du Japon en faveur de différents secteurs et pays.

La cohérence des politiques peut s'analyser d'un point de vue horizontal. Par exemple, dans le cadre de l'articulation entre action humanitaire, développement et recherche de la paix, le besoin de cohérence peut être particulièrement élevé, puisqu'un acteur donné peut mener des interventions relevant à la fois de la politique de développement, de la politique militaire et de la politique de sécurité. Dans le domaine de l'environnement également, il peut s'agir d'assurer la cohérence au niveau de l'articulation entre les secteurs de l'eau, de l'énergie et de l'alimentation ou entre les problématiques de l'égalité femmes-hommes et du changement climatique. Dans d'autres contextes, l'évaluation peut examiner la manière dont des domaines d'action autres que celui du développement – par exemple les échanges – influent sur l'intervention.

D'un point de vue vertical, la cohérence des politiques peut être appréhendée à des niveaux différents d'une même institution ou entre différentes composantes du système de financement du développement d'une même administration (par exemple, son organisme bilatéral, son institution de financement du développement (IFD) et son soutien multilatéral). Il peut aussi s'agir d'examiner dans quelle mesure l'intervention soutient ou au contraire compromet la réalisation des objectifs stratégiques à différents niveaux territoriaux – par exemple, d'évaluer le degré d'alignement d'une intervention locale à l'appui du développement sur les objectifs de développement nationaux et les interventions menées à l'échelle nationale, ou vice versa.

Cohérence externe

La cohérence externe recouvre deux dimensions principales : la concordance avec les engagements stratégiques extérieurs et la cohérence avec les interventions mises en œuvre par d'autres acteurs dans un contexte donné.

D'un point de vue stratégique, la cohérence externe concerne l'adéquation d'une intervention avec des engagements stratégiques extérieurs tels que les ODD et la manière dont ces engagements sont pris en compte dans la conception et la mise en œuvre de l'intervention. Il importe à ce stade de tenir compte de l'engagement d'une institution à atteindre les ODD, eu égard à la cible 14 de l'ODD 17, qui vise à « renforcer la cohérence des politiques de développement durable ». Il s'agit là d'un aspect important, qui permet de comprendre comment l'alignement stratégique et la redevabilité à l'égard des ODD sont pris en compte dans l'ensemble des activités et mis en œuvre dans la pratique.

Lorsqu'ils analysent la mise en œuvre d'une intervention dans un contexte donné, les évaluateurs devraient examiner la cohérence de l'intervention avec celles mises en œuvre par d'autres acteurs, en se posant par exemple la question de savoir comment les services sont fournis par divers acteurs et s'il existe des chevauchements ou au contraire des lacunes. La cohérence concerne la manière dont l'intervention apporte une valeur ajoutée par rapport à d'autres interventions et les mesures prises pour éviter les efforts redondants.

Toute évaluation de la cohérence externe devrait se concentrer sur l'intervention ou l'institution concernée, mais en la replaçant dans le contexte plus large des acteurs du domaine humanitaire et du développement durable. Il peut s'agir d'évaluer si les interventions sont conçues dans le cadre et en se servant des structures et des systèmes existants, par exemple les mécanismes de coordination au niveau d'un pays ou secteur.

Liens avec les autres critères

La cohérence est tout particulièrement liée à la pertinence, à l'efficacité et à l'impact.

- Alors que le critère de pertinence évalue l'intervention au niveau des besoins et des priorités des parties prenantes et des bénéficiaires directement concernés, le critère de cohérence se place au niveau supérieur et s'intéresse à l'adéquation de l'intervention au sein du système en général. L'un comme l'autre critère visent à déterminer la compatibilité de l'intervention avec le contexte, mais les points de vue adoptés sont différents.
- La cohérence représente souvent un angle d'approche utile pour engager un examen des effets non escomptés, qui peuvent ensuite être analysés au regard des critères d'efficacité et d'impact. Il peut arriver qu'une intervention atteigne ses objectifs (efficacité), mais que ses effets positifs soient annulés par d'autres interventions menées (en l'absence de toute cohérence) dans le même contexte.
- De même, il existe un lien entre cohérence et efficience : lorsque des interventions ne sont pas cohérentes, il peut y avoir des doublons, ce qui entraîne un gaspillage de ressources.

Intégrer la dimension de l'inclusion

La cohérence interne offre un angle d'analyse utile pour examiner la dimension de l'inclusion, en particulier en ce qu'elle se rapporte aux engagements, aux normes et aux critères relatifs aux droits humains. Les évaluateurs peuvent examiner la compatibilité de l'intervention avec les normes et les critères relatifs à l'inclusion et à l'égalité au niveau national ou institutionnel pour ce qui est des institutions de mise en œuvre, et du point de vue d'organisations locales comme les organisations de défense des droits des populations autochtones ou des personnes handicapées. L'évaluation de la cohérence peut apporter des éclairages utiles sur l'intérêt et la cohérence des activités visant à réduire l'exclusion, à atteindre les publics marginalisés et vulnérables et à réduire les inégalités femmes-hommes.

Les évaluateurs qui s'interrogent sur le degré d'inclusivité de l'impact et la pertinence de l'intervention devraient envisager d'analyser l'inclusion en lien avec la cohérence, car il existe des synergies entre ces trois dimensions de l'évaluation.

Difficultés posées par l'évaluation de la cohérence et solutions proposées

Le tableau ci-après décrit plusieurs des principales difficultés rencontrées lors de l'évaluation de la cohérence – liées notamment à l'étendue du champ d'évaluation, aux mandats et à la disponibilité des données – et propose à l'intention des évaluateurs comme des responsables d'évaluation des solutions pour les surmonter.

Tableau 4.2. Difficultés posées par l'évaluation de la cohérence

| Difficulté | Solutions : évaluateurs | Solutions : responsables d'évaluation |
|--|--|--|
| Dans beaucoup de contextes/pays, il est difficile d'évaluer la cohérence des politiques car, pour mener une évaluation rigoureuse, il faut avoir accès à une grande quantité de documents stratégiques, de données et d'interlocuteurs. | Ces difficultés ne peuvent être surmontées qu'à travers des évaluations qualitatives reposant sur les sources de données ou d'information existantes. Il importe également de mesurer l'évaluabilité au début d'une évaluation et de faire preuve d'ouverture et de transparence quant aux limites de l'évaluation. Lorsqu'il n'est pas possible d'accéder à certaines sources de données, il faut l'indiquer. Acquérir une connaissance du cadre législatif et des politiques de l'organisation et utiliser cette connaissance pour concevoir des plans appropriés de collecte et d'analyse des données primaires. | Les responsables d'évaluation devraient tenir compte de l'évaluabilité, des limites liées aux données et des problèmes d'accès lorsqu'ils commandent une évaluation. Ces aspects devraient être anticipés dans toute la mesure du possible et devraient, dès le début de l'évaluation, être examinés avec l'équipe chargée de l'évaluation en vue de décider de mesures d'atténuation. |
| Certaines organisations restreignent l'accès aux données en raison de la législation sur la protection des données et/ou de leurs politiques internes. | | Examiner les limites liées aux données et les problèmes d'accès au moment de la conception et de la commande d'une évaluation. Des échanges réguliers devraient être organisés avec l'équipe chargée de l'évaluation pour faciliter l'accès aux données et mettre au point des stratégies adaptées d'atténuation des risques. |
| L'évaluation de la cohérence entre tous les domaines de l'action publique pertinents peut considérablement élargir le champ de l'évaluation. | Collaborer avec le responsable d'évaluation pour que les priorités et le champ de l'évaluation soient clairement définis au début de l'évaluation. | Décider de la place que doit occuper la cohérence dans l'ordre des priorités en définissant le champ de l'évaluation et en hiérarchisant les domaines d'action à examiner. Ce choix devrait être fait en fonction de la finalité et de l'utilisation de l'évaluation. Dans le cadre de l'évaluation de la cohérence des politiques dans un secteur particulier, les institutions nationales devraient coordonner le plus étroitement possible cette évaluation au sein de ce secteur. |
| Les mandats de certains services d'évaluation peuvent être axés sur l'aide humanitaire et l'aide au développement seulement, ce qui peut limiter leur marge de manœuvre pour évaluer des questions relatives à la cohérence au sein de l'administration dans son ensemble. | Une attention particulière devrait être accordée à la phase de conception de l'évaluation afin de définir et d'affiner le champ d'évaluation, et de déterminer les modalités concrètes d'évaluation de la cohérence, conformément aux mandats respectifs des services d'évaluation – ou d'en indiquer les limites. | Pour évaluer la cohérence, il faut reconnaître que des champs d'action autres que le développement peuvent influencer sur une évaluation. En théorie, les évaluateurs sont ainsi plus libres de se pencher sur des domaines qui sortent du champ d'influence directe de l'intervention. En réalité, une telle liberté peut être difficile à concrétiser lorsqu'il s'agit d'évaluer des domaines qui n'entrent pas dans le cadre du mandat de l'institution. Il s'agit là d'une question d'évaluabilité qui devrait, dans l'idéal, être réglée avant l'évaluation. Les institutions devraient |

| Difficulté | Solutions : évaluateurs | Solutions : responsables d'évaluation |
|------------|----------------------------|--|
| | | déterminer s'il faut évaluer la cohérence dans des domaines de compétence réservés. S'il s'agit d'une priorité, les évaluateurs ou les responsables d'évaluation pourraient négocier avec d'autres services pour mener une évaluation conjointe. |

Exemples d'évaluation de la cohérence

Bien que la cohérence soit un nouveau critère pour le CAD de l'OCDE, elle fait partie des dimensions évaluées dans bon nombre d'évaluations réalisées au fil du temps. Le critère de cohérence est aussi couramment utilisé dans les évaluations de l'action humanitaire. Cette partie présente une sélection d'exemples d'évaluation de la cohérence tirés d'une évaluation stratégique de la cohérence des politiques au service du développement en Norvège, d'une évaluation de l'aide à la reconstruction après une catastrophe naturelle aux Philippines et d'une évaluation d'un portefeuille pays au Monténégro.

Encadré 4.4. Évaluation des efforts déployés par la Norvège pour garantir la cohérence des politiques au service du développement

Le gouvernement norvégien s'est engagé à plusieurs reprises à garantir la cohérence des politiques au service du développement dans le cadre de ses interventions. Par « cohérence des politiques » est entendu le fait d'intervenir sur des aspects plus larges du développement, en dehors de l'aide au développement – par exemple les échanges, les migrations, les investissements, le changement climatique et la sécurité. Cette évaluation examinait les « dilemmes » entre les politiques, c'est-à-dire les situations dans lesquelles il existe une contradiction manifeste entre les politiques de différents acteurs ou les domaines d'action. Le rapport d'évaluation plaide pour des échanges plus ouverts et plus transparents sur les dilemmes et priorités existant dans la réalité, compte tenu de l'existence de véritables dilemmes à résoudre et choix à effectuer.

L'évaluation avait pour but d'appréhender ces dilemmes et les difficultés rencontrées pour garantir la cohérence des politiques au service du développement. La cohérence était examinée à l'échelle d'un pays et à l'échelle mondiale et l'analyse portait aussi bien sur sa dimension interne qu'externe, à savoir :

- la cohérence interne des interventions pilotées par le ministère des Affaires étrangères et d'autres acteurs norvégiens ;
- la cohérence externe, pour déterminer le degré d'alignement des objectifs stratégiques des acteurs sur les objectifs de développement internationaux.

Étudiant le cas du Myanmar, l'évaluation examinait les incohérences entre la politique de développement et les autres domaines de l'action publique. Le Myanmar avait été retenu en raison du large éventail d'acteurs du développement norvégiens, traditionnels ou non, qui y interviennent et des conflits qui peuvent apparaître entre leurs priorités stratégiques respectives au niveau local. L'évaluation adoptait une méthode intéressante pour appréhender les dilemmes : elle examinait les aspects au sujet desquels les organisations non gouvernementales (ONG), les organisations politiques ou les institutions internationales avaient signalé l'existence d'incohérences entre les priorités et interventions de la Norvège. Deux dilemmes ont ainsi été mis en relief : l'engagement des entreprises norvégiennes au Myanmar et l'engagement du pays en faveur de la paix, d'une part, et l'implication dans le développement des ressources nationales et l'engagement en faveur de la paix, d'autre part. L'analyse a aussi mis en lumière une difficulté majeure, tant en termes d'évaluation de la cohérence que d'adoption de mesures visant à la renforcer, à savoir que le ministère n'est pas mandaté pour

intervenir au sujet de décisions stratégiques prises par d'autres ministères, si bien qu'il lui est plus difficile de comprendre et d'éliminer les incohérences. Dans le cadre de l'évaluation de la cohérence à l'échelle mondiale, la cohérence externe avec les engagements mondiaux de la Norvège était également examinée et l'évaluation a relevé des difficultés à aligner la politique intérieure sur les priorités internationales en matière de développement.

Cette évaluation montre ce que peut être une application réfléchie du critère de cohérence en illustrant les difficultés rencontrées pour mettre en œuvre des interventions cohérentes qui tiennent compte à la fois du contexte de mise en œuvre et des priorités institutionnelles.

Source : Norad (2018^[7]), *Evaluation of Norwegian Efforts to Ensure Policy Coherence for Development*, <https://www.norad.no/contentassets/4ac3de36fbd4229811a423f4b00acf7/8.18-evaluation-of-norwegian-efforts-to-ensure-policy-coherence-for-development.pdf>

Encadré 4.5. Évaluation de la cohérence de l'aide à la reconstruction après une catastrophe naturelle aux Philippines

Cette évaluation portait sur les programmes canadiens d'aide internationale et sur le rôle joué par le Canada en tant que donneur dans les situations de catastrophe naturelle. Elle montre comment la cohérence d'une intervention peut être considérée d'un point de vue externe comme interne. Elle examinait le programme de reconstruction de 20.5 millions de dollars canadiens mis en œuvre à la suite du typhon Haiyan ainsi que la relation bilatérale entre le Canada et les Philippines entre 2013-14 et 2018-19.

Évaluation de la cohérence

Même si elle a été réalisée avant l'adoption par le CAD de l'OCDE du nouveau critère d'évaluation de la cohérence, cette évaluation mettait clairement l'accent sur cet aspect, sans toutefois le préciser expressément. Elle en analysait les dimensions tant externe qu'interne.

- Cohérence externe – l'évaluation a examiné le rôle joué par le Canada en tant que donneur, son comportement dans les différents contextes opérationnels et politiques et l'adéquation de l'intervention avec ces contextes. Elle a mis en lumière une concordance avec les priorités et le plan national de relance du gouvernement philippin et a évalué la complémentarité avec d'autres programmes en place dans la région.
- Cohérence interne – l'évaluation a examiné le programme d'aide internationale du Canada et l'a replacé dans le contexte plus large d'autres programmes d'aide internationale. Elle s'est intéressée en particulier à l'alignement du programme sur la politique globale du Canada et a évalué dans quelle mesure l'intervention incarnait l'approche du pays en matière de secours en cas de catastrophe.

L'évaluation a aussi montré combien la cohérence et la pertinence se rejoignent, l'intervention étant constamment mise en lien avec les besoins et les priorités des parties prenantes et du contexte – deux éléments également pris en compte pour analyser la pertinence. Toutefois, elle est allée encore plus loin en ce qu'elle a tenté d'appréhender non seulement la capacité d'adaptation de l'intervention mais aussi le degré d'alignement et les considérations prises en compte dans le processus d'élaboration et de mise en œuvre d'une intervention claire, cohérente, qui corresponde aux intérêts plus larges du Canada aux Philippines et soit cohérente avec les autres programmes de secours en place.

Source : Affaires mondiales Canada (2019^[8]), *Évaluation de l'aide à la reconstruction en cas de catastrophe naturelle aux Philippines de 2013-14 à 2018-19* <https://www.international.gc.ca/gac-amc/publications/evaluation/2019/endra-earcn-philippines.aspx?lang=fra>

Encadré 4.6. Évaluer la cohérence du programme de coopération pour le développement de la Slovénie avec le Monténégro

Cette évaluation décrivait la mise en œuvre d'activités de coopération pour le développement entre la République de Slovénie et le Monténégro au cours de la période 2013-16. Elle avait pour but de fournir des données factuelles sur les résultats obtenus (c'est-à-dire, dans quelle mesure le programme atteint les objectifs de la coopération pour le développement), d'analyser les raisons susceptibles d'expliquer les bons ou les mauvais résultats et de formuler des recommandations pour l'élaboration de politiques, programmes et projets futurs.

La cohérence était l'un des huit critères évalués (figuraient également parmi ces derniers la gestion du programme, la mise en œuvre et la valeur ajoutée pour la Slovénie) et l'évaluation visait à répondre à la question suivante :

- Des contradictions avec d'autres politiques ont-elles entravé la mise en œuvre et empêché d'atteindre les objectifs de développement ou, au contraire, les politiques se renforcent-elles mutuellement ?

L'évaluation portait également sur la cohérence externe avec d'autres politiques liées à la coopération pour le développement, notamment le processus de négociation en vue de l'adhésion à l'Union européenne et les investissements dans les infrastructures environnementales. Elle a conclu que les programmes étaient en adéquation avec les autres politiques.

S'intéressant aussi à la complémentarité avec les activités d'autres donateurs, elle a constaté que le manque de systèmes de coordination des donateurs au Monténégro avait nui à la complémentarité et empêché d'éviter les doublons.

Source : Ministère des Affaires étrangères slovène (2017^[9]), *Evaluation of Slovenia's Development Cooperation with Montenegro 2013-2016 period: Final Report*, <https://www.gov.si/assets/ministrstva/MZZ/Dokumenti/multilateral/razvojno-sodelovanje/Development-cooperation-with-Montenegro-evaluation-final-report.pdf>

Efficacité

Définition de l'efficacité :

Efficacité : l'intervention atteint-elle ses objectifs ?

Mesure dans laquelle les objectifs et les résultats de l'intervention ont été atteints, ou sont en train de l'être, y compris les résultats différenciés entre populations.

Remarque : l'analyse de l'efficacité suppose de prendre en compte l'importance relative des objectifs ou des résultats. Le terme « efficacité » est également utilisé comme indicateur global de la mesure dans laquelle une intervention a eu ou devrait avoir un impact pertinent et durable, avec efficacité et cohérence.

Qu'est-ce que l'efficacité et pourquoi est-elle importante ?

L'efficacité permet de comprendre dans quelle mesure une intervention atteint ou est en passe d'atteindre ses objectifs. Ce critère peut être un moyen de déterminer si une intervention a donné les résultats prévus, par quel processus ces résultats ont été obtenus et quels ont été les facteurs décisifs à cet égard ainsi que les éventuels effets non escomptés. L'efficacité porte sur les résultats les plus étroitement attribuables à l'intervention et doit donc être distinguée de l'impact, qui examine les effets plus généraux et les changements de portée plus vaste.

Pour évaluer la réalisation des objectifs le long de la chaîne de résultats/causalité, il faut bien comprendre les buts et objectifs de l'intervention. L'application du prisme de l'efficacité peut donc aider les évaluateurs, les responsables ou administrateurs de programmes et d'autres acteurs à définir (ou évaluer) des objectifs clairs. De même, le critère d'efficacité peut aider les évaluateurs à déterminer si l'obtention (ou l'absence) de résultats s'explique par des faiblesses dans la mise en œuvre ou la conception de l'intervention.

Au titre de ce critère, les évaluateurs devraient également déterminer les effets non escomptés. Dans l'idéal, les responsables de projet auront déjà identifié les risques au stade de la conception de l'intervention, ce dont pourront se servir les évaluateurs lorsqu'ils entament leur évaluation. Il importe d'examiner les effets non escomptés pour repérer les effets tant négatifs (par exemple, l'accentuation d'une dynamique de conflit) que positifs (par exemple, des innovations qui améliorent l'efficacité). Les institutions qui commandent des évaluations peuvent fournir aux évaluateurs des orientations concernant les normes minimales à appliquer pour identifier les effets non escomptés, en particulier quand ces derniers revêtent la forme de violations de droits humains ou d'autres conséquences non intentionnelles graves.

La définition de l'efficacité encourage les évaluateurs et les responsables à se poser des questions importantes sur la répartition des résultats entre différents groupes, que cette répartition soit intentionnelle ou non. Il s'agit d'un choix délibéré, qui vise à renforcer la prise en compte de l'équité, conformément à la priorité stratégique des ODD consistant à « ne laisser personne de côté ». Les évaluateurs sont donc encouragés à examiner les aspects relatifs à l'équité, que celle-ci constitue ou non un objectif spécifique de l'intervention. Ce type d'analyse nécessite des données et exige parfois d'investir des ressources — ce qui est souvent justifié étant donné les précieux éclairages que l'évaluation peut apporter³.

Pour tirer des conclusions sur l'efficacité d'une intervention, les évaluateurs devraient se concentrer sur les résultats les plus importants dans le contexte en question et pour les destinataires de l'évaluation. À cet égard, l'expression « importance relative » insiste sur l'idée que l'évaluateur devrait opérer un jugement évaluatif et soulever l'importance des objectifs et des résultats atteints ou non atteints/attendus, y compris les conséquences non intentionnelles.

Comprendre l'efficacité : éléments d'analyse

La définition de l'efficacité est sous-tendue par les principales notions suivantes : réalisation des objectifs, importance variable des objectifs et des résultats, résultats différenciés selon les groupes et compréhension des facteurs qui influent sur les réalisations.

Réalisation des objectifs

L'évaluation de l'efficacité vise en premier lieu à établir si une intervention a atteint les résultats escomptés à différents niveaux de la chaîne de résultats (en général, au niveau des extrants et des réalisations, mais parfois aussi de l'impact). Cette chaîne devrait être définie au moment de la conception de l'intervention et constitue la principale référence pour la gestion, le suivi et l'évaluation.

Il est très difficile d'évaluer l'efficacité d'une activité si les objectifs énoncés ou les résultats prévus sont formulés de manière vague ou ambiguë ou ont évolué au fil de l'intervention, sans être actualisés ni restructurés. Les responsables de l'intervention devraient au minimum expliquer pourquoi les objectifs ont évolué et quels sont les nouveaux objectifs. À défaut, les évaluateurs devront consulter les documents relatifs à l'intervention ou interroger les parties prenantes pour reconstituer la logique sous-jacente aux évolutions successives de l'intervention. Sur la base de cette reconstitution, ils pourront alors déterminer le niveau de pertinence et de réalisation effective des nouveaux objectifs.

L'évaluation de l'efficacité revêt également de l'importance dans le cadre des programmes adaptatifs, auxquels des modifications sont apportées selon un processus itératif, à partir des retours d'information des parties prenantes, des premiers résultats et de l'évolution du contexte. Dans ce type de programmes, la conception et la mise en œuvre d'une intervention peuvent connaître de multiples changements progressifs survenant au fil du temps. Il importe dans ce cas que les évaluateurs examinent et analysent les théories du changement au regard des systèmes plus larges dans lesquels s'inscrit l'intervention et tiennent compte de tout document indiquant comment et pourquoi des changements ont été apportés. L'évaluation de l'efficacité devrait prendre en compte les objectifs actuels. Elle devrait s'appuyer sur un examen de la logique et de la nécessité sous-tendant les changements apportés en matière de mise en œuvre ou d'objectifs (qui sont souvent mises en évidence dans les versions actualisées de la théorie du changement ou du cadre de résultats).

L'évaluation de l'efficacité peut consister à établir quels sont les changements observables au sein du groupe cible ou du contexte tout au long de la mise en œuvre de l'intervention, ainsi que la causalité des changements observés à différents niveaux – c'est-à-dire montrer que les changements ont été induits par l'intervention ou que celle-ci y a contribué plutôt que d'autres facteurs liés à l'environnement ou, à défaut, une autre intervention. Les méthodologies devraient être conçues pour permettre à l'évaluateur de déterminer comment les résultats ont été obtenus et pourquoi (facteurs explicatifs) ils ont été atteints, insuffisants ou nuls.

L'évaluation de l'efficacité consiste également à examiner les résultats de l'intervention. Selon la définition du CAD de l'OCDE, les résultats englobent les extrants, les réalisations ou les impacts (escomptés ou non, positifs et/ou négatifs) de l'intervention (OCDE, 2002_[10]). Par conséquent, l'évaluation de l'efficacité peut également impliquer d'évaluer d'éventuels effets non escomptés, positifs ou négatifs, induits par l'intervention. La mise en œuvre d'une intervention peut toujours avoir des effets sociaux, économiques ou environnementaux non escomptés ou des effets qui sont inattendus mais auraient pu être prévus. L'évaluation devrait donc veiller à prendre en compte les effets non mentionnés dans les objectifs de l'intervention, ce qui peut supposer d'aller jusqu'à examiner les avantages ou les risques qui pourraient découler de ces effets non escomptés (prévisibles ou non). L'évaluation de l'efficacité consiste également à déterminer dans quelle mesure l'intervention a contribué à la réalisation d'objectifs de développement nationaux ou d'autres objectifs de développement pertinents dans le contexte – en ce qui concerne la contribution potentielle de l'intervention, c'est un aspect qui relève de la pertinence.

Évaluer l'importance relative des résultats obtenus

Lorsqu'ils évaluent l'efficacité, les évaluateurs s'intéressent à la réalisation (ou non) des divers objectifs et résultats de l'intervention. Si certains des objectifs – mais pas tous – ont été atteints, ils examineront leur importance relative pour pouvoir tirer des conclusions sur l'efficacité de l'intervention. Cet examen peut s'appuyer sur l'analyse de la pertinence, qui devrait avoir permis d'évaluer les éventuelles différences entre les priorités et les besoins des diverses parties prenantes. Les évaluateurs peuvent ainsi en conclure que l'intervention a été efficace sur certains plans mais pas sur d'autres, ou qu'elle l'a été du point de vue de certaines parties prenantes, mais pas d'autres.

Résultats différenciés

Les évaluateurs devraient s'intéresser à la dimension de l'inclusivité et de l'équité des résultats entre les différents groupes de bénéficiaires – qu'il s'agisse de personnes ou d'institutions. L'appréhension des résultats différenciés peut nécessiter d'examiner dans quelle mesure la conception et la mise en œuvre de l'intervention étaient inclusives. À titre d'illustration, l'évaluation peut examiner le processus suivi pour formuler les objectifs de l'intervention, ce qui consiste notamment à déterminer si les objectifs ont été formulés sur la base d'une analyse des besoins et d'un processus de consultation des diverses parties prenantes (dont le principal groupe cible). Ce processus peut aussi être un moyen de comprendre si l'intervention a laissé passer des occasions d'obtenir des résultats pour la population ou les bénéficiaires ciblés, notamment des occasions de contribuer à un changement à plus long terme, par exemple un recul des inégalités. Les évaluateurs peuvent examiner les résultats non escomptés ou inattendus ainsi que les résultats escomptés en tenant compte du fait que certains changements et transformations s'opèrent de manière subtile et/ou sur le long terme et peuvent être difficiles à quantifier.

Facteurs exerçant une influence

L'analyse des facteurs qui influent sur les résultats est importante parce qu'elle aide les évaluateurs à comprendre *pourquoi* une intervention a ou non atteint ses objectifs, ce qui permet aux partenaires de repérer les domaines à améliorer. Les facteurs peuvent être internes ou externes à l'intervention. Ils peuvent être liés à la gestion, aux ressources humaines, aux aspects financiers ou réglementaires, à des modifications dans la mise en œuvre ou à des écarts par rapport aux plans. La qualité de la mise en œuvre (et le respect des protocoles de mise en œuvre) est souvent décisive pour l'efficacité. Elle devrait donc être décrite avant d'évaluer l'efficacité, l'efficience, l'impact et la viabilité/durabilité.

S'agissant des facteurs externes, les évaluateurs devraient examiner les effets positifs et négatifs induits par le contexte de l'intervention, ceux-ci contribuant à la réalisation (ou non) des résultats. Ce processus peut notamment consister à évaluer la capacité d'adaptation de l'intervention face aux évolutions du contexte. Les évaluations peuvent aussi examiner l'opportunité dans le temps des résultats (par exemple, l'échelonnement du soutien aux groupes ou environnements ciblés, contribuant à l'obtention des résultats).

Liens avec les autres critères

L'efficacité est liée aux autres critères, en particulier à la pertinence et à l'impact :

- La définition des objectifs relève de la pertinence et la réalisation de ces objectifs dépend de l'efficacité. Une intervention non pertinente peut fort bien être mise en œuvre avec efficacité. En pareil cas, les évaluateurs devront faire preuve de jugement au moment de rédiger leurs conclusions générales, les constatations relatives à un critère ne pouvant être ignorées au profit de celles portant sur d'autres critères.
- L'efficacité et l'impact sont des critères complémentaires qui mettent l'accent sur des niveaux différents de la chaîne de résultats. L'efficacité a trait à l'obtention des résultats par rapport aux

objectifs d'une intervention, c'est-à-dire au niveau des extrants et des réalisations, alors que l'impact concerne les résultats de plus vaste portée, plus précisément les résultats généraux énoncés et les facteurs qui contribuent à leur obtention. D'une manière générale, les responsables de l'intervention et les évaluateurs devraient veiller à ce qu'une distinction claire soit faite entre les différents niveaux de résultats (intrants, extrants, réalisations et impact) et à ce que les aspects à évaluer dans le cadre de chaque critère soient clairement définis.

Intégrer la dimension de l'inclusion

La définition de l'efficacité encourage l'examen approfondi de la question de l'équité entre différents groupes. Les évaluateurs devraient apprécier le degré d'inclusion de différents groupes de bénéficiaires dans l'intervention et la mesure dans laquelle des principes fondamentaux comme l'équité, la non-discrimination et la redevabilité ont été pris en compte à tous les stades de l'intervention, depuis la conception jusqu'aux résultats. Conformément à l'objectif consistant à ne « laisser personne de côté », la question de savoir dans quelle mesure l'intervention répondait aux besoins des plus marginalisés devrait faire l'objet d'une attention particulière. Il importe d'examiner l'obtention et la répartition des résultats pour différents groupes de bénéficiaires et d'expliquer les écarts éventuels.

De plus, les évaluateurs devraient se pencher sur la question de savoir si, comment et pourquoi les résultats contribuent à réduire les inégalités. Au titre du critère d'efficacité, les évaluateurs devraient examiner l'effet de certaines activités sur le bien-être de groupes spécifiques et déterminer si ces activités fournissent aux participants des moyens d'autonomisation.

Difficultés posées par l'évaluation de l'efficacité et solutions proposées

Le tableau ci-après décrit les difficultés liées aux données, à la formulation des objectifs et à l'attribution des résultats et propose à l'intention des évaluateurs comme des responsables d'évaluation des solutions pour les surmonter.

Tableau 4.3. Difficultés posées par l'évaluation de l'efficacité

| Difficulté | Solutions : évaluateurs | Solutions : responsables d'évaluation |
|--|---|---|
| Absence de données appropriées, informatives ou ventilées. | Le manque de données (par exemple l'absence d'évaluations de référence ou à mi-parcours) peut entraver l'évaluation de la réalisation des objectifs, les effets de causalité étant plus difficiles à attribuer et l'origine et la nature des contributions étant plus difficiles à déterminer. Il est possible de reconstituer des données de référence, mais les données ainsi obtenues sont en général moins précises et moins fiables que des données de référence dûment établies. Parmi les autres méthodes que les évaluateurs peuvent adopter figurent : 1) la triangulation et la validation des données disponibles ; 2) l'utilisation d'autres méthodes ; et 3) la construction d'un récit sur la contribution apportée par l'intervention. | Dans l'idéal, la planification en matière de disponibilité des données devrait être effectuée au stade de la conception de manière à garantir la collecte de données appropriées, informatives et ventilées tout au long de la durée de vie de l'intervention. Les responsables d'évaluation devraient avoir une vision claire des données disponibles et des autres sources pouvant être utilisées, et devraient coopérer avec les évaluateurs pour mettre au point d'autres approches lorsqu'aucune donnée n'est disponible. Utiliser les diagnostics posés dans d'autres évaluations et les rapports de recherche dans le domaine concerné. Examiner les enseignements tirés de travaux antérieurs sur la définition d'indicateurs dans des domaines de travail connexes, par exemple les cibles des ODD, ainsi que les moyens permettant de s'en inspirer ou de les appliquer. |
| Les objectifs/énoncés d'objet sont mal formulés, vagues ou difficiles à mesurer. | Au stade initial d'une évaluation, veiller à ce que les objectifs de l'intervention soient vérifiables, réalistes et conformes aux normes internationales applicables aux interventions de développement. Les indicateurs utilisés pour mesurer la réalisation des objectifs devraient eux aussi être validés au regard de critères généralement admis comme les critères SMART (spécifique, | Travailler en coopération avec les évaluateurs afin de définir des critères clairs permettant de mesurer la réalisation des objectifs conformément aux normes généralement admises (par exemple, les indicateurs SMART). Pour des évaluations tenant compte de la problématique femmes-hommes, envisager de passer des indicateurs SMART et RACER (pertinent, accepté, crédible, simple et robuste) aux |

| Difficulté | Solutions : évaluateurs | Solutions : responsables d'évaluation |
|----------------------------|--|--|
| | mesurable, atteignable, réaliste et défini dans le temps). Dans l'idéal, il faudrait que ce processus procède d'une démarche participative associant à la fois les évaluateurs et les responsables d'évaluation. | indicateurs CREAM (clair, pertinent, économique, adéquat et évaluable) et SPICED (subjectif, participatif, interprété, comparé et recoupé, responsabilisant, divers et désagrégé). Améliorer la définition des objectifs et des résultats et les rendre plus concrets, dans le cadre de la théorie (du changement) du programme définie à l'époque. En l'absence de théorie du changement, tenter d'en reconstruire une valable pour contextualiser la logique interne de l'intervention. Améliorer <i>a posteriori</i> la cohérence de la formulation en remplaçant le projet dans son contexte à partir de pratiques participatives et en associant les parties prenantes. |
| Attribution des résultats. | La question de l'attribution de la causalité ou de la détermination de la contribution d'une intervention fait partie des difficultés auxquelles se heurte l'évaluation de l'efficacité, en particulier à long terme. La construction d'un récit sur la contribution apportée par l'intervention et la mise au point d'une triangulation appropriée des sources de données peuvent aider l'évaluateur à déterminer si les résultats peuvent être attribués à l'intervention évaluée. | Les opinions, les évaluations et les jugements des parties prenantes qui connaissent bien l'intervention/l'entité et pourraient être des partenaires dans sa mise en œuvre ou dans l'obtention des résultats à l'appui du développement sont pris en compte. |

Exemples d'évaluation de l'efficacité

Cette partie présente une sélection d'exemples d'évaluation de l'efficacité tirés d'une évaluation d'un programme d'assistance électorale et d'une évaluation d'un programme-pays au Cabo Verde.

Encadré 4.7. Évaluation de l'efficacité et de l'inclusivité du programme australien d'assistance électorale

Depuis les années 70, l'Australie fournit une assistance électorale à divers pays en développement afin de les accompagner dans leur transition démocratique et dans l'instauration de démocraties formelles qui reposent sur des bases solides et fonctionnent. Cette évaluation portait sur l'aide apportée par l'Australie à l'organisation d'élections nationales dans huit pays de la région Asie-Pacifique entre 2006 et 2016. Elle fournit un bon exemple d'application du critère d'efficacité. Elle prend en particulier en compte l'évolution des objectifs au fil du temps et examine également l'efficacité sous l'angle de l'inclusion, de la problématique femmes-hommes et des groupes vulnérables.

L'évaluation commençait par examiner les objectifs globaux du programme australien d'assistance électorale, avant de s'intéresser plus précisément à l'efficacité et à l'inclusion dans trois grands domaines : le renforcement des systèmes de gestion électorale, la promotion de la participation et le soutien à l'organisation d'élections.

Elle soulignait également que les objectifs de l'Australie concernant l'assistance électorale avaient évolué pendant la phase de mise en œuvre, ce dont tenait compte l'analyse. Par ailleurs, l'évaluation examinait également de manière rigoureuse les risques associés à la fourniture d'une assistance électorale effective et inclusive, en tenant compte des facteurs externes susceptibles de limiter l'efficacité des interventions actuelles et futures.

Dans le cadre de l'analyse des objectifs, l'évaluation se concentrait particulièrement sur le caractère inclusif des résultats, ce qui illustre utilement la manière dont on peut examiner des résultats différenciés selon les groupes de population pour évaluer l'efficacité. L'obtention des résultats était examinée à travers un prisme sexospécifique, l'évaluation s'appuyant sur des constatations qualitatives

pour montrer que les normes sociales et les obstacles juridiques et institutionnels se traduisaient par une participation électorale beaucoup plus faible pour les femmes.

L'évaluation tenait aussi compte de l'inclusion des groupes vulnérables, plus précisément des personnes handicapées, et constatait que des barrières culturelles contribuaient également à limiter l'équité des résultats entre différents groupes malgré les actions menées pour améliorer l'accessibilité physique. Forte de ces constatations, l'évaluation a formulé des recommandations pour les interventions futures, axées sur l'amélioration de l'équité des résultats en tant que moyen de favoriser la réalisation plus large des résultats de l'intervention et des objectifs de développement.

Source : Arghiros et al. (2017^[11]), *Making it Count: Lessons from Australian Electoral Assistance 2006-16*, <http://www.oecd.org/derec/australia/australia-electoral-assistance-2006-2016.pdf>

Encadré 4.8. Évaluer l'efficacité de la stratégie et du programme-pays de la Banque africaine de développement (BAfD) au Cap Vert

L'évaluation du soutien apporté par la BAfD au secteur de l'électricité au Cap Vert fournit un exemple d'application du critère d'efficacité dans la pratique, en tenant compte de la pertinence de la conception et de la mise en œuvre et en accordant une attention particulière aux objectifs liés à l'égalité entre les sexes et à l'inclusion sociale. La BAfD poursuivait plusieurs objectifs généraux à travers son programme, à savoir, notamment, aider le pays à favoriser une croissance inclusive grâce au développement des infrastructures et à des réformes de la gouvernance, et – il s'agissait là d'un objectif stratégique – accompagner le pays dans sa transition vers une « croissance verte » ou un développement bas carbone et résilient face au changement climatique.

Pour évaluer l'efficacité de l'intervention et déterminer si ces objectifs avaient été atteints, l'évaluation entreprenait de déterminer si l'intervention avait débouché sur les résultats attendus prévus, soit :

- l'amélioration de l'accès de la population à l'électricité ;
- l'accroissement de la proportion des énergies renouvelables ;
- l'augmentation de la superficie de terres arables et de l'approvisionnement en eau ;
- l'amélioration de la gouvernance.

Dans ce cadre, l'évaluation analysait également les facteurs qui avaient influé sur l'intervention, qu'il s'agisse de facteurs internes, comme la solidité de la conception et les progrès dans la mise en œuvre, ou de facteurs externes. Une attention particulière était accordée à l'égalité entre les sexes et à l'inclusion sociale dans le cadre de l'analyse de l'équité des résultats.

Si l'évaluation a conclu que de solides avancées avaient été réalisées s'agissant d'aider le pays à favoriser une croissance inclusive grâce au développement des infrastructures et aux réformes de la gouvernance, elle a constaté que les ressources et les instruments disponibles étaient insuffisants pour atteindre l'objectif stratégique visant à accompagner le pays dans sa transition vers une « croissance verte » ou un développement bas carbone et résilient face au changement climatique. Il en est ressorti que l'examen de l'efficacité peut renseigner sur la réalisation ou non des résultats prévus au titre d'une intervention, ainsi que sur les facteurs qui ont influé sur cette réalisation.

Source : IDEV (2018^[12]), *Cabo Verde: Evaluation of the Bank's Country Strategy and Program 2008–2017: Summary Report*, <https://idev.afdb.org/sites/default/files/Evaluations/2020-03/Executive%20Summary%20Cabo%20Verde%20CSPE%20%28Fr%29.pdf>

Efficiency

Définition de l'efficacité :

Efficacité : les ressources sont-elles utilisées de manière optimale ?

Mesure dans laquelle l'intervention produit, ou est susceptible de produire, des résultats de façon économique et dans les temps.

Remarque : le terme « économique » désigne la conversion des intrants (fonds, expertise, ressources naturelles, temps, etc.) en extrants, réalisations et impacts de la façon la plus économiquement avantageuse possible, par rapport aux options envisageables dans le contexte. L'expression « dans les temps » désigne le fait de respecter les délais fixés ou des délais raisonnablement adaptés aux exigences du contexte en évolution. Il peut s'agir d'évaluer l'efficacité opérationnelle (mesure dans laquelle l'intervention a été bien gérée).

Qu'est-ce que l'efficacité et pourquoi est-elle importante ?

Ce critère permet de vérifier si les ressources allouées à une intervention peuvent être justifiées par ses résultats, ce qui est essentiel d'un point de vue pratique comme politique. L'efficacité est un critère important pour beaucoup de catégories de parties prenantes, notamment les pouvoirs publics, la société civile et les bénéficiaires. Une meilleure utilisation de ressources limitées permet d'obtenir plus de résultats dans le cadre de la coopération pour le développement, par exemple sur le plan de la réalisation des ODD, domaine dans lequel les besoins sont énormes. L'efficacité intéresse particulièrement les pouvoirs publics, qui ont des comptes à rendre à leurs contribuables, lesquels remettent souvent en cause le rapport coût-efficacité de divers programmes et politiques, en particulier les décisions relatives à la coopération internationale pour le développement, qui font souvent l'objet d'une surveillance plus étroite.

L'efficacité est également importante d'un point de vue opérationnel. Beaucoup d'interventions se heurtent à des problèmes de faisabilité et de mise en œuvre, en particulier en ce qui concerne la manière dont les ressources sont utilisées. Parce qu'elle rend les responsables redevables de la manière dont ils prennent les décisions et gèrent les risques, l'évaluation de l'efficacité contribue à les inciter à veiller à la bonne conduite des programmes.

Plusieurs hypothèses et points importants doivent être rappelés :

- La notion de ressources s'entend au sens large et recouvre l'ensemble des coûts économiques (humains, environnementaux, financiers et en termes de temps). Elle ne doit pas être confondue avec le budget du programme ou les fonds dépensés.
- De même, les résultats s'entendent au sens large et recouvrent l'ensemble de la chaîne de résultats, à savoir les extrants, les réalisations et les impacts. Selon le type d'évaluation, certaines organisations associent l'efficacité aux seuls extrants. Dans le présent document toutefois, ce critère est défini et conceptualisé de manière à encourager l'évaluation de l'efficacité en lien également avec des effets de portée plus vaste comme l'impact, même si c'est là un exercice qui s'avère souvent difficile.
- Évaluabilité : la capacité d'évaluer l'efficacité, l'impact, la cohérence et la viabilité/durabilité influe sur les conclusions relatives à l'efficacité.
- L'efficacité renvoie à un arbitrage entre diverses solutions possibles pour parvenir aux mêmes résultats, dans les limites des ressources disponibles. Pour pouvoir comparer leur rapport coût-

efficacité, il faut identifier des solutions qui peuvent réellement être mises en œuvre et soient comparables en termes de qualité et de résultats.

C'est pourquoi l'analyse de l'efficacité doit résolument se faire sur fond d'analyse du contexte : en effet, dans un contexte donné, il peut être plus coûteux, mais aussi plus important et justifiable en termes d'impact sur le développement, d'atteindre les bénéficiaires cibles.

Comprendre l'efficacité : éléments d'analyse

Évaluer l'efficacité suppose d'analyser les dimensions principales que sont l'efficacité économique, l'efficacité opérationnelle et l'opportunité dans le temps.

Efficacité économique

C'est le principal élément à prendre en compte dans l'analyse de l'efficacité. L'efficacité économique désigne ici l'absence de gaspillage et la conversion des intrants en résultats de la façon la plus économiquement avantageuse possible. Il s'agit notamment d'évaluer l'efficacité des résultats à tous les niveaux de la chaîne de résultats – extrants, réalisations et impacts –, ce qui suppose d'évaluer dans quelle mesure les choix appropriés ont été faits et des arbitrages opérés au stade de la conception et durant la mise en œuvre. Ces choix concernent notamment la répartition des ressources entre les différents groupes cibles et les différentes périodes, ainsi que les options disponibles pour acquérir les intrants aux conditions du marché.

Efficacité opérationnelle

L'efficacité opérationnelle est également un aspect important à prendre en compte. Il s'agit de savoir dans quelle mesure les ressources sont utilisées de manière optimale pendant la mise en œuvre. Les questions qui peuvent être posées pour l'analyser sont notamment les suivantes : les ressources humaines et financières ont-elles été utilisées conformément à ce qui était prévu, pleinement et de manière appropriée (ou y a-t-il eu une mauvaise allocation des ressources, une sous-utilisation ou au contraire un dépassement du budget) ? L'affectation des ressources a-t-elle été modifiée en fonction de l'évolution des besoins ? Les risques ont-ils été gérés ? Des décisions de nature à améliorer l'efficacité ont-elles été prises en réponse à de nouvelles informations ? Les décisions prises en matière de logistique et d'achats étaient-elles optimales ?

Opportunité dans le temps

Étroitement liée à l'efficacité économique et à l'efficacité opérationnelle, cette dimension suppose d'abord de se demander si et dans quelle mesure les résultats ont été obtenus dans les délais prévus. C'est aussi l'occasion de vérifier si le calendrier était dès le départ réaliste ou approprié, et s'il a été adapté de manière raisonnable au cours de l'intervention, étant donné la fréquence, dans de nombreuses interventions, des facteurs externes et des changements de programme. Les évaluateurs doivent apprécier si, au fil de l'évolution de la situation, des mesures ont été prises pour surmonter les obstacles et limiter les retards dans la gestion de l'intervention.

Liens avec les autres critères

Comme souligné précédemment, les différents critères sont liés les uns aux autres et devraient être considérés comme autant de prismes à travers lesquels analyser l'intervention plutôt que comme des ensembles rigides. Entre autres liens d'interdépendance entre l'efficacité et d'autres critères figurent les suivants :

- Pertinence et efficience : un aspect essentiel de la pertinence opérationnelle concerne le fait de savoir si la conception de l'intervention était bien adaptée au contexte, ouvrant la voie aux considérations de faisabilité et de capacités. Concrètement, la question de savoir si la conception de l'intervention était réaliste et pouvait être mise en œuvre influe également directement sur l'efficience. Ainsi, pour cet aspect particulier, l'évaluateur peut se retrouver à examiner ces deux critères simultanément.
- Efficience et résultats : l'efficience supposant d'évaluer dans quelle mesure les ressources utilisées ont été converties en résultats, tous les aspects des résultats (c'est-à-dire, les questions qui émergent au fil de l'évaluation de l'efficacité, de l'impact et de la viabilité/durabilité) devraient être examinés. L'efficience opérationnelle est étroitement liée à l'efficacité et à l'impact. L'analyse du bon fonctionnement d'une intervention suppose souvent d'examiner simultanément l'efficacité et l'efficience. Ce constat vaut en particulier lorsqu'on détermine, par exemple, où se situent les goulets d'étranglement et comment y remédier, ou lorsqu'on veut s'assurer que les ressources sont affectées là où elles sont nécessaires.

Intégrer la dimension de l'inclusion

Le critère d'efficience permet aussi aux évaluateurs de comprendre comment la dimension de l'inclusion est intégrée et appréhendée dans la gestion de l'intervention et dans quelle mesure l'utilisation des ressources reflète les différences d'expérience et de résultats entre les publics. L'obtention des résultats a un coût variable selon les bénéficiaires, les « plus défavorisés » étant les plus difficiles à toucher – et ceux pour lesquels le coût est le plus élevé. L'analyse de l'efficience devrait donc reposer sur une bonne compréhension des inégalités et des dynamiques de pouvoir dans le contexte concerné, ainsi que de la manière dont l'intervention répond à la nécessité de favoriser des changements porteurs de transformation pour lutter contre les inégalités profondes. C'est une occasion essentielle d'examiner si un engagement à l'égard du programme d'action visant à « ne laisser personne de côté » (et de l'objectif du Programme 2030 consistant à opérer un changement transformateur pour les groupes marginalisés) a été dûment et effectivement mis en œuvre aux niveaux de la gestion, de la prise de décision et de l'affectation des ressources.

À cet égard, l'analyse peut notamment viser à déterminer comment et pourquoi les ressources sont réparties entre les différents groupes ciblés par une intervention, et dans quelle mesure cette répartition est fondée sur les besoins des groupes marginalisés et le dialogue engagé avec eux. Les évaluateurs peuvent se pencher sur la question de savoir si des résultats inclusifs et équitables ont été obtenus moyennant un coût raisonnable, comment la notion de « coût raisonnable » est définie et déterminée et quelles sont les variations de coût d'un groupe de bénéficiaires à l'autre. Par exemple, si l'intervention a vocation à toucher des groupes spécifiques, il conviendra de se poser la question de savoir si les ressources allouées sont suffisantes et justifiées pour atteindre ce but.

Un aspect essentiel consiste à comprendre quels sont les groupes qui parviennent à se faire entendre et dont le point de vue est pris en compte au moment de la prise de décisions sur la conception des politiques, l'affectation des fonds et les personnes chargées de contrôler et de superviser ces processus. Lorsque la logique d'une intervention et les plans y afférents visent notamment à modifier des structures et une dynamique de pouvoir inégalitaires, l'évaluation devrait examiner dans quelle mesure ce but a été atteint ou si l'intervention a sans le vouloir renforcé des structures et une dynamique inégalitaires préexistantes.

Il importe également d'évaluer si, dans le cadre des interventions, des données de suivi pertinentes et ventilées sont recueillies de manière à faciliter la prise de décisions pertinentes par les acteurs chargés de leur mise en œuvre concernant l'axe principal des activités/objectifs et les ressources à allouer au développement inclusif.

Difficultés posées par l'évaluation de l'efficience et solutions proposées

L'application adéquate du critère d'efficience dépend entièrement de la nature de l'intervention elle-même et varie selon les projets, les programmes, les programmes-pays et les politiques. L'exemple ci-dessous, qui provient du ministère des Affaires étrangères néerlandais, illustre la manière dont ce critère a été appliqué pour évaluer la politique dans le secteur de l'eau, où l'efficience est en grande partie appréhendée au regard de la coordination et d'aspects pratiques de la planification et des partenariats au sein d'un réseau complexe de relations (Encadré 4.11).

Le Tableau 4.4 décrit plusieurs des principales difficultés rencontrées lors de l'évaluation de l'efficience et propose à l'intention des évaluateurs comme des responsables d'évaluation des solutions pour les surmonter.

Une des décisions fondamentales consiste à se demander si l'efficience doit être évaluée à l'aide des indicateurs économiques et instruments connexes traditionnels (analyse coût-avantages, taux de rendement, analyse coût-efficacité, évaluation comparative, etc.) et comment utiliser ces outils⁴. La réponse à cette question dépend de la finalité de l'évaluation, de l'intervention et des résultats visés, de la faisabilité, des données/ressources disponibles et du public cible.

L'utilité de différents outils d'analyse dépend également de l'approche retenue aux stades de la conception/l'approbation au sein de l'institution concernée, cette approche ayant d'importantes conséquences sur la disponibilité des informations nécessaires pour entreprendre différents types d'analyse. Au sein des banques multilatérales de développement et dans le cadre de certains programmes publics d'investissement en capital, il existe des orientations très claires sur l'évaluation économique, sociale et environnementale *ex ante* et, de plus en plus, sur la conduite d'une analyse de ces projets et programmes en tenant compte des disparités entre les sexes. Pour évaluer l'efficience au stade de l'examen préalable ou de l'approbation du projet (en général *ex ante*, c'est-à-dire avant la mise en œuvre de l'intervention), il est logique de recourir aux mêmes outils que ceux utilisés pour l'évaluation elle-même. Les règles et les recommandations d'action adoptées par l'institution déterminent également en partie quelles données et quels indicateurs sont accessibles aux évaluateurs (par exemple, si les taux de rendement ont été estimés pendant l'évaluation économique, le cas échéant, et si d'autres interventions ont été recensées).

Tableau 4.4. Difficultés posées par l'évaluation de l'efficience

| Difficulté | Solutions proposées : évaluateurs | Solutions proposées : responsables d'évaluation |
|--|---|--|
| Trouver des bases de comparaison adaptées | Les évaluateurs ont généralement du mal à trouver des bases de comparaison parfaites et utilisent donc ce qu'ils ont à leur disposition. Ils doivent néanmoins faire preuve de prudence lorsqu'ils effectuent des comparaisons et devraient effectuer une analyse de sensibilité. À titre d'exemple, l'efficience d'une intervention mise en œuvre dans des situations de fragilité ou de conflit devrait uniquement être comparée avec celle d'une intervention mise en œuvre dans un contexte présentant le même niveau de défis. | Lorsque les responsables d'évaluation ont accès à des données ou des analyses susceptibles de fournir des informations de référence utiles aux évaluateurs, prévoir d'inclure ces ressources dans les documents communiqués aux évaluateurs. |
| S'y retrouver dans les diverses notions de l'efficience ainsi que dans les méthodes et outils associés | Les travaux consacrés à l'efficience en tant que concept économique sont nombreux. Les évaluateurs peuvent les consulter pour déterminer quelles méthodes et quels outils sont les plus appropriés compte tenu de l'évaluation à mener. Le document de travail du BMZ (Palenberg, 2011 ^[13]) en dresse une excellente synthèse. | |

| Difficulté | Solutions proposées : évaluateurs | Solutions proposées : responsables d'évaluation |
|---|---|---|
| Manque d'expertise en matière d'analyse de l'efficacité | Veiller à ce qu'il existe une compréhension de ce qui est nécessaire pour réaliser une analyse de l'efficacité et de la manière dont cette analyse peut être intégrée dans le processus d'évaluation (en termes d'approche comme d'expertise). | Veiller à ce que les termes de référence des évaluations précisent le degré d'expertise requis en matière d'analyse de l'efficacité. |
| Données et mesure | Les données sur les avantages, les résultats et les coûts peuvent être difficiles à obtenir. Ainsi, les évaluations d'impact peuvent fournir des estimations quantitatives des changements intervenus au niveau des réalisations sous l'effet de l'intervention, mais, d'une manière générale, ne pas rendre compte des coûts estimés et encore moins de l'ensemble des coûts. | Veiller à ce que les évaluateurs aient accès aux données nécessaires pour entreprendre une analyse de l'efficacité (par exemple, des données sur le rapport coût-avantages d'une intervention ou action donnée, en termes de budget, de ressources humaines, de temps, etc.). |
| Considérations temporelles | Pour l'évaluation économique comme pour l'évaluation de l'efficacité, le choix entre les périodes au cours desquelles les ressources sont utilisées et celles où les résultats sont obtenus est crucial. Les résultats et l'ensemble des coûts (par exemple, les « coûts environnementaux ») peuvent se concrétiser sur de nombreuses années, si bien que l'efficacité réelle au moment où le programme est mené ou l'évaluation effectuée ne donne pas une image complète de la situation. | L'efficacité est influencée par une série de choix opérés pendant la mise en œuvre du programme ou de l'intervention et notamment par la manière dont les responsables ont pris des mesures d'adaptation face à la réorientation des priorités, ont anticipé les risques et les ont gérés, ont traité les dépassements de coûts, etc. |

Exemples d'évaluation de l'efficacité

Cette partie présente une sélection d'exemples d'évaluation de l'efficacité en relation avec un programme d'électrification rurale, des subventions aux intrants agricoles et un vaste portefeuille d'activités portant sur la gestion de l'eau.

Encadré 4.9. Évaluation de l'efficacité du soutien de la Banque mondiale en faveur de l'électrification rurale

Cette étude portait sur un programme de prêts de la Banque mondiale visant à soutenir l'électrification rurale dans un ensemble de pays entre 1980 et 2006. Elle analysait notamment un portefeuille de 120 projets, dix études de cas nationales et des données d'enquêtes disponibles (enquêtes sur l'énergie et enquêtes auprès des ménages concernant 12 pays).

Puisant dans les vastes ensembles de données de cette grande banque multilatérale de développement dotée d'une longue expérience dans le secteur concerné, l'étude examinait l'efficacité sous différents angles, comparant notamment les coûts et les avantages. L'analyse du portefeuille de projets s'appuyait sur les estimations des taux de rendement des projets, qui figurent habituellement dans les rapports d'achèvement validés des projets d'infrastructure. En outre, les effets sur le bien-être étaient estimés sur la base d'estimations du surplus du consommateur, qui permettent aux évaluateurs de déterminer les bienfaits des interventions et les coûts associés. Cette évaluation offre une illustration d'une approche permettant d'évaluer l'impact, les aspects relatifs à l'équité (toucher les populations pauvres) et les arbitrages entre viabilité, efficacité économique et allocative, impact global et impact par groupe cible.

Source : IEG, (2008^[14]), *The Welfare Impact of Rural Electrification: A Reassessment of the Costs and Benefits* ; <https://openknowledge.worldbank.org/handle/10986/6519>.

Encadré 4.10. Évaluer l'efficacité des subventions aux intrants agricoles en Afrique subsaharienne

L'étude dont il est question avait pour but d'évaluer les subventions aux intrants agricoles au Ghana, au Malawi, en Tanzanie et en Zambie, notamment les subventions dites « intelligentes », qui sont spécifiquement conçues pour obtenir un maximum d'effets à moindre coût.

Évaluation de l'efficacité

Du fait de la nature du sujet, l'efficacité économique – soit, dans le cas présent, la question de savoir si le coût des subventions était, à différents égards, justifié – constituait une question centrale.

L'étude reposait sur des estimations de la variation de la productivité agricole et examinait si cette variation était à la hauteur des coûts, c'est-à-dire si le rendement était satisfaisant, et quels facteurs – par exemple le ciblage – jouaient un rôle. Les observations de l'étude concernant la Zambie, par exemple, sont les suivantes :

L'utilisation accrue d'intrants agricoles semble avoir des effets substantiels sur la production de maïs, mais cette augmentation de la production a un coût très élevé. D'après les meilleures estimations disponibles, importer du maïs pour couvrir les besoins en milieu urbain coûterait moins cher qu'accroître la production en Zambie au moyen du programme pour la sécurité alimentaire (ZFSP) dans sa forme actuelle. Les subventions aux intrants pourraient néanmoins être justifiées si le manque d'efficacité manifeste était compensé par des considérations d'équité ou par un développement durable à plus long terme du secteur des intrants. Or, rien n'est plus incertain.

Ainsi, tout en portant sur l'efficacité, l'étude prend également spontanément en compte plusieurs autres critères, dont certaines dimensions de la durabilité/viabilité (par exemple, quelle stratégie de sortie est appliquée) et de l'impact/efficacité des subventions, ainsi que l'impact global.

Source : Baltzer et Hansen (2011^[15]), *Agricultural input subsidies in Sub-Saharan Africa: Evaluation Study*, <http://www.oecd.org/derec/49231998.pdf>.

Encadré 4.11. Évaluer l'efficacité au niveau des politiques : la politique néerlandaise d'aide au développement visant à améliorer la gestion de l'eau

Cet examen portait sur la politique d'aide au développement mise en œuvre par le ministère néerlandais des Affaires étrangères entre 2006 et 2016 pour améliorer la gestion intégrée des ressources en eau – une priorité d'action de plus en plus importante pour les Pays-Bas. Dans ce cadre avaient été établis divers partenariats avec des organisations multilatérales, des gouvernements nationaux, des ONG et le secteur privé. Les deux principaux critères retenus pour les besoins de l'évaluation étaient l'efficacité des politiques ainsi que leur efficacité.

Évaluation de l'efficacité

La question d'évaluation posée pour examiner cet aspect était libellée ainsi :

- La mise en œuvre de la politique était-elle organisée et assurée de manière à contribuer à la réalisation des principaux résultats visés concernant les groupes d'utilisateurs de l'eau, la qualité technique et l'entretien des infrastructures physiques ?

Pour répondre à cette question, les évaluateurs ont examiné dans quelle mesure la réalisation des principaux résultats s'écartait (en termes de coûts et de délais) de ce qui était prévu, et pour quelles raisons. Ils ont examiné les résultats à chaque niveau de la chaîne de résultats, notamment la qualité, les réalisations et la viabilité/durabilité. L'évaluation a par exemple constaté que la planification de la gestion participative de l'irrigation subissait parfois des retards et des dépassements de coûts parce que la complexité sociale et technique des problématiques locales exacerbait les facteurs politiques et institutionnels, ce qui contribuait à une exécution plus lente et à des dépassements de coûts.

Source : IOB (2017^[16]), *Tackling major water challenges: Policy review of Dutch development aid policy for improved water management*, <https://english.iob-evaluatie.nl/publications/policy-review/2017/12/01/418-%E2%80%93-iob-%E2%80%93-policy-review-of-dutch-development-aid-policy-for-improved-water-management-2006-2016-%E2%80%93-tackling-major-water-challenges>.

Impact

Définition de l'impact :

Impact : Quelle différence l'intervention fait-elle ?

Mesure dans laquelle l'intervention a produit, ou devrait produire, des effets importants et de vaste portée, positifs ou négatifs, intentionnels ou non.

Remarque : le critère d'impact porte sur l'importance ultime et les effets potentiellement transformateurs de l'intervention. Il vise à déterminer quels sont ses effets sociaux, environnementaux et économiques à plus long terme ou à plus vaste échelle que ceux déjà évalués au titre du critère d'efficacité. Au-delà des résultats immédiats, ce critère vise à évaluer les conséquences indirectes, secondaires et potentielles de l'intervention, à travers l'examen des évolutions globales et durables des systèmes ou des normes, ainsi que des effets potentiels sur le bien-être des populations, les droits humains, l'égalité des sexes et l'environnement.

Qu'est-ce que l'impact et pourquoi est-il important ?

Le critère d'impact invite à s'interroger sur la différence que fait réellement l'intervention. C'est au titre de ce critère que les évaluateurs examinent les effets que l'intervention a, *in fine*, sur le plan du développement – si elle a induit des transformations qui comptent réellement pour les populations. C'est l'occasion d'adopter une perspective plus large et une vision globale. En effet, il est facile de se laisser absorber par les aspects liés aux activités courantes d'une intervention et de se contenter de suivre le cadre de référence auquel se conforment ceux qui la mettent en œuvre. Le critère d'impact incite les évaluateurs à aller plus loin et à déterminer quelles évolutions ont été opérées, au bénéfice de quelles populations. Le Manuel d'évaluation de l'Agence suédoise de coopération pour le développement international (Asdi) souligne l'importance de ce point (Molund et Schill, 2004_[17]), indiquant que :

Le critère d'impact apporte un correctif important à ce qui pourrait sinon être une préoccupation excessivement étroite concernant les intentions de ceux qui planifient et gèrent les interventions de développement et à l'indifférence qui pourrait en résulter vis-à-vis des points de vue des groupes cibles et d'autres parties prenantes clés.

Même si le terme « impact » est couramment employé, il importe de noter qu'il prête souvent à confusion, ce qui peut influencer sur la façon dont les parties prenantes appréhendent l'évaluation. En premier lieu, dans un contexte politique, il peut être utilisé de façon vague pour désigner les « résultats » au sens large, englobant à la fois l'efficacité et l'impact tels que définis dans le présent document, mais aussi d'autres aspects de la performance. En deuxième lieu, depuis quelques années, il est souvent confondu avec l'expression « évaluation d'impact », qui renvoie à des méthodologies spécifiques permettant d'établir des liens de causalité statistiquement significatifs entre l'intervention et les effets observés⁵. Dans cette acception, l'impact peut désigner des résultats à n'importe quel stade de la chaîne des résultats, y compris les produits, et renvoie presque toujours à des effets *souhaités*. C'est pourquoi il importe de clarifier dès le départ avec les parties prenantes la façon dont elles comprennent ce terme et d'expliquer comment il est utilisé dans le contexte de l'évaluation comme un critère permettant d'examiner les effets de vaste portée.

Parmi les questions susceptibles d'être posées au titre du critère d'impact figurent les suivantes :

- L'intervention a-t-elle induit un changement significatif dans la vie des bénéficiaires visés ?
- Comment l'intervention a-t-elle produit des effets de vaste portée (par exemple, des évolutions au niveau des normes ou des systèmes) ?
- L'intervention a-t-elle apporté les mêmes bénéfices à tous les groupes cibles visés, y compris les plus défavorisés et les plus vulnérables ?
- L'intervention est-elle porteuse de transformations : induit-elle des changements durables au niveau des normes – y compris celles régissant les relations entre les sexes – et des systèmes, intentionnellement ou non ?
- L'intervention conduit-elle à d'autres changements, notamment des résultats susceptibles d'être portés à plus grande échelle ou reproduits ?
- Comment l'intervention contribuera-t-elle à faire évoluer la société pour le mieux ?

Comprendre l'impact : éléments d'analyse

La définition de l'impact englobe les notions clés que sont les effets de vaste portée, l'importance, les effets différenciés, les effets non intentionnels et les changements porteurs de transformation.

Importance

Le critère d'impact permet de répondre à la question de savoir quelle différence a réellement fait l'intervention. Il examine l'importance de l'intervention et ses résultats de plus vaste portée, c'est-à-dire la mesure dans laquelle l'intervention a fait une différence pour les personnes concernées.

Cette définition vise à encourager les évaluateurs à examiner l'intervention sous différents angles, en fonction du contexte. Les évaluateurs devraient dûment prendre en compte le contexte, les besoins et les priorités des bénéficiaires ciblés par l'intervention, les objectifs stratégiques convenus par les institutions concernées et la nature de l'intervention elle-même. Cet élément d'analyse peut également être appliqué lors de l'examen des résultats non escomptés d'une intervention.

Lorsqu'ils évaluent l'importance⁶, les évaluateurs devraient être conscients de la nécessité d'adopter différents points de vue ainsi qu'une approche systématique fondée sur les besoins des parties prenantes. Ils devraient également prendre des mesures pour éviter que leurs préjugés (implicites) et leurs jugements de valeur n'influencent sur leur évaluation de l'importance de l'intervention.

Impact différencié

Conformément à l'engagement des ODD de « ne laisser personne de côté » et de protéger les droits humains, y compris l'égalité entre les sexes, il importe d'évaluer les impacts différenciés. Les impacts positifs globaux peuvent masquer d'importants effets de répartition négatifs. Il est essentiel d'en tenir compte au stade de la conception de l'évaluation, voire au stade de la conception de l'intervention, afin de permettre le suivi, puis l'évaluation, de l'impact par groupe cible. Cet objectif suppose une planification précoce lors de la conception et de l'évaluation, afin de s'assurer que les évaluateurs disposent, dans la mesure du possible, de données désagrégées, et peut également nécessiter l'examen d'une série de paramètres relatifs à l'exclusion et à l'inclusion. Si des données désagrégées sont disponibles, une analyse granulaire sera nécessaire.

Effets non intentionnels

Les évaluateurs devraient déterminer si une intervention a des effets non intentionnels. Cette analyse devrait notamment viser à établir dans quelle mesure les impacts étaient prévus ou envisagés au moment

de la conception de l'intervention. Les effets non intentionnels peuvent être positifs ou négatifs. Lorsqu'ils sont positifs, les évaluateurs devraient examiner leur importance globale et les possibilités d'innovation ou bien d'expansion ou de reproduction de cet impact positif sur d'autres interventions. Les évaluateurs devraient accorder une attention particulière aux effets négatifs, notamment ceux qui sont susceptibles d'être significatifs, y compris – mais sans s'y limiter – l'impact sur l'environnement ou les effets non intentionnels sur des groupes vulnérables.

Changements porteurs de transformation

Au sens de la définition, les changements porteurs de transformation sont des « évolutions globales et durables des systèmes ou des normes », qui peuvent être appréhendées comme s'attaquant aux causes profondes, ou aux facteurs systémiques, de la pauvreté, des inégalités, de l'exclusion et de la dégradation de l'environnement, et dont le Programme de développement durable à l'horizon 2030 juge qu'elles sont nécessaires pour atteindre les ODD. Il est de plus en plus courant que les interventions visent à contribuer à un changement transformationnel et, lorsqu'ils évaluent l'impact, les évaluateurs sont de plus en plus appelés à répondre à des questions relatives aux effets sur les normes et les systèmes (systèmes sociaux, économiques ou politiques). Par exemple, une évaluation peut examiner dans quelle mesure les rôles traditionnels des femmes et des hommes ont été modifiés d'une manière ou d'une autre (voir Encadré 4.12).

Liens avec les autres critères

À l'instar des autres critères, celui relatif à l'impact interagit, sur le plan conceptuel, avec d'autres critères :

- **Impact et efficacité** : l'impact comme l'efficacité ont trait aux résultats obtenus par l'intervention, aux effets sur les bénéficiaires et aux résultats non escomptés. La différence entre les deux critères dépend en grande partie de la manière dont l'intervention et ses objectifs ont été conçus. L'efficacité concerne généralement la réalisation des objectifs fixés (à n'importe quel niveau de la chaîne de résultats), tandis que l'impact concerne toujours les effets de plus vaste portée, qui ne seraient sinon pas automatiquement pris en compte (parce que n'ayant pas été inclus dans les objectifs). Une autre façon d'envisager la distinction entre les deux critères consiste à considérer le fait qu'avec le temps, de nombreuses interventions peuvent être jugées efficaces, sans toutefois « s'additionner » pour induire le changement de plus vaste portée ou porteur de transformation souhaité. Si l'impact n'est pas évalué, ces tendances ne sont pas analysées. La formulation des résultats escomptés d'une intervention sera différente selon les institutions, qui appliquent souvent des critères différents pour définir la chaîne de résultats. Ainsi, dans certaines institutions, les interventions doivent chacune être reliées à des objectifs plus généraux ; dans d'autres, les effets immédiats sont les seuls pris en compte. Pour appliquer les deux critères, les institutions devront veiller à ce que les exigences en matière de conception des interventions soient claires et cohérentes. Lorsque des interventions de moindre envergure, comme des projets, ne précisent pas systématiquement les liens existant avec des objectifs généraux, il importe que la politique d'évaluation prévoie une analyse d'impact – à défaut, les effets de vaste portée ne seront pas pris en compte (pour une illustration, voir Encadré 4.12).
- **Impact et cohérence** : l'analyse de l'impact supposant l'adoption d'une perspective globale, ce critère s'accorde naturellement bien avec les considérations relatives à la cohérence, étant donné que les effets d'une intervention dépendent presque toujours d'autres interventions, d'objectifs stratégiques et d'arbitrages, ainsi que des systèmes dans lesquels l'intervention s'inscrit. Citons à titre d'exemple les programmes de coopération pour le développement qui soutiennent le renforcement des systèmes de santé publique dans les pays en développement, où l'impact est déterminé non seulement par le programme en question, mais aussi par les politiques nationales

ou mondiales en matière de tarification et de réglementation des produits pharmaceutiques ou de recrutement des personnels de santé.

- Impact et efficacité : comme indiqué dans la section relative au critère d'efficacité, les évaluateurs qui s'intéressent à l'efficacité au sens large doivent examiner les résultats obtenus (par exemple, l'impact, la durabilité) dans leur globalité et comparer les résultats avec les ressources.
- Impact et viabilité/durabilité : les deux critères s'intéressent, dans une certaine mesure, à la pérennité des résultats. L'impact met l'accent sur la dimension temporelle en examinant les changements porteurs de transformation (qui sont durables par nature). La viabilité/durabilité examine si les bénéfices de l'intervention perdureront. En tant que critère, la viabilité/durabilité couvre un champ plus large, car elle prend en compte les conditions permettant d'assurer la durabilité et les liens entre la durabilité économique, sociale et environnementale d'une intervention.

Intégrer la dimension de l'inclusion

Les notions de changements porteurs de transformation, d'impact différencié et d'importance sont toutes intrinsèquement liées à celle d'inclusion. Il importe ici de comprendre la nature de l'impact qui s'est produit et qui en a bénéficié. Cet impact a-t-il contribué de manière significative à la transformation des systèmes d'oppression et pourrait-il de ce fait conduire à un changement durable pour les groupes marginalisés et vulnérables ? Les évaluateurs devraient chercher à comprendre dans quelle mesure ces effets non intentionnels découlent des inégalités structurelles existant au sein de systèmes plus larges et de l'impact des interventions sur ces systèmes.

La définition révisée de l'impact met l'accent sur les résultats généraux d'une intervention, y compris les effets sociaux et économiques à long terme. Elle englobe les changements porteurs de transformation au niveau des systèmes et des normes, y compris « les effets potentiels sur le bien-être des populations, les droits humains, l'égalité des sexes et l'environnement ». L'analyse de l'impact permet aux évaluateurs d'avoir une vision plus globale de la manière dont une intervention contribue aux changements porteurs de transformation, à l'équité, aux droits humains et à l'autonomisation, et les renforce.

Difficultés posées par l'évaluation de l'impact et solutions proposées

Des six critères, l'impact est celui qui est souvent le plus difficile à évaluer et à appréhender. Ci-après figurent quatre des principaux obstacles à surmonter, assortis de propositions sur la manière de procéder.

En règle générale, l'évaluation de l'impact nécessite davantage de ressources et nettement plus de données primaires que celle des autres critères, et devrait uniquement être prévue au moment de la conception d'une évaluation lorsque ces ressources sont susceptibles d'être disponibles ou peuvent être obtenues à cette fin. D'un autre côté, c'est souvent le domaine qui suscite le plus d'attention de la part des utilisateurs, étant donné qu'il s'agit de déterminer si, et comment, l'intervention a changé le cours des choses. Par conséquent, l'investissement en temps et en effort nécessaire pour intégrer l'analyse de cette dimension dans l'évaluation est souvent justifié.

Une autre difficulté liée à l'évaluation de l'impact concerne les délais fixés par les institutions commanditaires. En effet, ceux-ci correspondent souvent à la date de clôture de l'intervention. Si cette exigence était assouplie, l'impact pourrait être examiné sur une plus longue période de temps, ce qui permettrait de mieux appréhender les effets dont l'existence n'apparaît qu'après l'achèvement d'une intervention.

Le tableau ci-après décrit plusieurs des principales difficultés rencontrées lors de l'évaluation de l'impact et propose à l'intention des évaluateurs comme des responsables d'évaluation des solutions pour les surmonter.

Tableau 4.5. Difficultés posées par l'évaluation de l'impact

| Difficulté | Solutions : évaluateurs | Solutions : responsables d'évaluation |
|---|---|---|
| Clarté des explications relatives à la signification de la notion d'impact pour les parties prenantes concernées et les utilisateurs de l'évaluation. | Convenir dès le départ de la définition appliquée, qu'il s'agisse de la terminologie du CAD de l'OCDE ou d'une autre notion, et s'y conformer systématiquement. Vérifier régulièrement, au cours des discussions avec les parties prenantes, si leur interprétation du terme a changé. | |
| Clarté des explications relatives à l'impact attendu de l'intervention et à la manière dont il sera atteint. | Cet aspect dépend de la conception initiale de l'intervention et de la disponibilité d'un cadre de résultats clair, assorti d'une théorie du changement. À défaut d'un tel cadre, il sera difficile de procéder à une évaluation de l'impact fondée sur une théorie (ce qui est hautement souhaitable) et cette théorie devra donc être reconstruite en faisant référence à des groupes cibles spécifiques. Certaines interventions ont tendance à viser des objectifs trop généraux et démesurés par rapport à ce que l'intervention peut réaliser. | |
| Déterminer à quel degré l'intervention produit l'impact. | <p>Déterminer les voies par lesquelles l'intervention a été à l'origine de l'impact ou y a contribué, le cas échéant, depuis l'intervention jusqu'aux résultats, lorsque ces voies risquent d'être moins faciles à retracer/explicites à des niveaux supérieurs. Le recours à des méthodes comme l'analyse de contribution peut être utile à cet égard⁷.</p> <p>En s'appuyant sur des orientations/méthodologies spécifiques portant sur la manière de mesurer l'impact, étudier les possibilités et opérer des choix réfléchis et réalistes à propos des méthodes en fonction de l'objectif de l'évaluation, du contexte, des questions auxquelles l'évaluation doit répondre, de la disponibilité des données, ainsi que des ressources disponibles pour l'évaluation et de la possibilité de les mobiliser. Une question essentielle consiste à savoir si les méthodes permettront uniquement d'évaluer l'ampleur de l'impact de l'intervention en question ou si elles permettront également de déterminer pourquoi et comment il s'est produit.</p> | Vérifier que les ressources disponibles pour l'évaluation et les compétences des personnes mobilisées (s'agissant de l'application des méthodes) sont suffisantes pour évaluer l'impact. |
| Disponibilité des données, notamment des bases de référence et des indicateurs. | Pour évaluer les effets de plus vaste portée, la disponibilité des données peut s'avérer particulièrement problématique, notamment le manque de bases de référence et d'indicateurs. Il s'agit en partie d'une question d'évaluabilité qui devrait être prise en compte dès le début de l'évaluation, mais aussi d'un aspect qui éclaire le choix des méthodes appropriées pour exploiter au mieux les différents types de données qualitatives et quantitatives. | Établir la disponibilité des données au cours du processus de rédaction des termes de référence de l'évaluation. |
| L'intervention a d'importants effets inattendus ou non intentionnels. | Il est souvent difficile de déterminer les résultats inattendus ou non intentionnels d'une intervention, car ils sortent souvent du champ des cadres de suivi et de la collecte de données. Pour cerner ce qui importe le plus aux bénéficiaires et aux autres personnes touchées par l'intervention (positivement ou négativement), les évaluateurs devraient se montrer disposés à explorer d'autres domaines que la seule documentation officielle du projet et les données attendues. L'importance relative de ces effets sera probablement différente selon les parties prenantes. | Offrir aux évaluateurs un espace où ils puissent écouter les personnes touchées par l'intervention (positivement ou négativement) et les aider à attirer l'attention sur les effets inattendus ayant une importance à leurs yeux. |

Exemples d'évaluation de l'impact

Cette partie présente une sélection d'exemples d'évaluation de l'impact en lien avec des interventions portant sur l'autonomisation des familles, sur la consolidation de la paix, sur le soutien à la recherche, sur l'aménagement du territoire et sur la violence à l'égard des femmes et des filles.

Encadré 4.12. Évaluer l'efficacité et l'impact du programme « Plus de familles en action » en Colombie

Une évaluation de l'impact d'un point de vue opérationnel et de l'impact sur le bien-être du programme « Plus de familles en action », mis en œuvre en Colombie, offre un exemple intéressant de la façon dont des évolutions profondes et globales peuvent être appréhendées et recensées, notamment la manière dont certains aspects sont évalués au titre de l'efficacité et de l'impact, sous des angles légèrement différents. Cette évaluation examinait le rôle des *Madres Lideres*, des femmes qui servent de porte-parole de leur communauté et qui ont été formées pour aider à la mise en œuvre du programme.

- D'un point de vue opérationnel, l'évaluation analysait le rôle joué par les *Madres Lideres* pour soutenir l'efficacité du programme, mettant en lumière le rôle essentiel des femmes dans sa mise en œuvre et ses résultats.
- Au niveau plus général, l'évaluation a déterminé que leur participation au programme avait eu des effets positifs sur l'autonomisation des femmes, ainsi que des effets potentiellement porteurs de transformation sur les communautés dans leur ensemble en modifiant la perception du rôle des femmes dans la société. Ces évolutions générales des normes sont appréhendées au titre du critère d'impact.

Source : Economía Urbana et Ipsos (2019^[18]), *Evaluación de operaciones del programa "más familias en acción" y de resultados del componente de bienestar comunitario [Évaluation du programme « Plus de familles en action »]*, https://colaboracion.dnp.gov.co/CDT/Sinergia/Documentos/Evaluacion_MFA_Informe_Resultados.pdf

Encadré 4.13. Évaluer l'impact du soutien apporté par l'Asdi à la consolidation de la paix dans des pays touchés par un conflit ou sortant d'un conflit

Cette évaluation constitue un exemple intéressant de la prise en compte du critère d'impact pour examiner des interventions complexes menées sur une longue période dans des situations de fragilité. Elle évaluait l'impact en termes d'incidence globale des efforts déployés par l'Agence suédoise de coopération pour le développement international (Asdi) à l'appui de la consolidation de la paix et examinait les résultats directs, indirects, intentionnels et non intentionnels, positifs et négatifs.

L'évaluation appréhendait les impacts à long terme en adoptant une perspective à très long terme, couvrant les activités de consolidation de la paix menées en Somalie, au Rwanda, en Bosnie et au Guatemala du début des années 90 à la fin des années 2010. Outre l'impact global, elle s'intéressait particulièrement aux femmes et aux groupes marginalisés. La méthodologie du « changement le plus significatif » était utilisée pour évaluer la contribution apportée par l'intervention.

Source : Bryld (2019^[19]), *Evaluation of Sida's Support to Peacebuilding in Conflict and Post-Conflict Contexts: Somalia Country Report*, https://publikationer.sida.se/contentassets/1396a7eb4f934e6b88e491e665cf57c1/eva2019_5_62214en.pdf

Encadré 4.14. Évaluation de l'impact dans le cadre d'une évaluation stratégique de l'action publique

Dans le cadre d'une évaluation du soutien apporté et des recherches menées par 21 autorités danoises, engagées dans 40 projets dans 18 pays partenaires, une des principales conclusions dressées au titre du critère d'impact concernait l'importance des effets globaux sur l'autonomisation des institutions partenaires. L'évaluation concluait qu'une des principales réalisations de l'intervention avait été de donner des moyens d'action aux autorités partenaires au-delà des objectifs officiels du projet. Cette conclusion renvoie à la capacité du Danemark d'interagir et de nouer des liens avec les principaux acteurs du secteur public, du secteur privé, de la société civile et du développement, et d'influencer les résultats plus généraux des politiques et des programmes. Ce type de réalisations apportent des bénéfices mutuels aux deux partenaires, mais passent le plus souvent « inaperçues » dans le suivi des interventions. Or, il s'agit d'un résultat essentiel du programme, qui doit entrer en ligne de compte pour en apprécier toute la valeur ajoutée.

Source : PEM Consult (2020^[20]), *Evaluation of the Danish Strategic Sector Cooperation*, https://um.dk/en/danida-en/results/eval/eval_reports/publicationdisplaypage/?publicationID=CBE77158-1D4D-46E3-A81F-9B918218FAFF

Encadré 4.15. Évaluation de l'impact, de la diffusion et de la montée en puissance d'une approche globale de l'aménagement du territoire aux Philippines

Examinant l'impact d'une intervention d'aménagement du territoire, cette évaluation analysait ses effets à moyen et long terme, sur la base d'indicateurs environnementaux et socioéconomiques.

Elle appliquait une approche théorique et un mélange de techniques, s'appuyant sur des données de panel issues d'une enquête à plusieurs niveaux, des entretiens qualitatifs, des groupes de discussion, une revue de la littérature, une analyse des documents d'aménagement du territoire et des données et des informations géographiques. Une reconstruction complète de la théorie du changement de l'intervention a été effectuée par les évaluateurs.

L'évaluation s'articulait autour d'une analyse d'impact rigoureuse qui mesurait et quantifiait les effets dans cinq domaines allant de l'amélioration des structures administratives et des conditions dans les administrations en charge de la planification à l'amélioration du bien-être de la population concernée, en passant par la gestion des ressources naturelles, les mesures et les activités liées à la gestion des risques de catastrophe et le fonctionnement de la gouvernance locale.

Les données de panel étaient constituées de 3 000 ménages, répartis dans 300 *barangays* (unités administratives locales) dans 100 municipalités, au sein de 11 provinces de la région des Visayas, et deux mesures avaient été réalisées dans le temps (2012 et 2016). Parmi ces ménages, certains avaient reçu une aide de l'agence de développement allemande, la GIZ (*Deutsche Gesellschaft für Internationale Zusammenarbeit*), d'autres non. La méthode d'évaluation de l'impact intégrait également un modèle quasi-expérimental fondé sur l'appariement des scores de propension afin d'identifier des « jumeaux statistiques », et se basait sur plusieurs dizaines de caractéristiques des municipalités, des *barangays* et des ménages.

Source : Leppert et al. (2018^[21]), *Impact, Diffusion and Scaling-Up of a Comprehensive Land-Use Planning Approach in the Philippines: From Development Cooperation to National Policies*, https://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/Zusammenfassung_Deutsch_%20Eval-2018_Philippinen_final_web-2.pdf

Encadré 4.16. Évaluation de l'impact du programme « Girl Empower » au Libéria

Cette évaluation, réalisée par le Comité international de secours, fournit un exemple spécifique d'évaluation de l'impact d'une intervention multidimensionnelle visant à lutter contre la violence à l'égard des femmes et des filles : le projet « Girl Empower – Mentoring and Cash Transfer Intervention to Promote Adolescent Wellbeing in Liberia ». Cette intervention a été mise en œuvre dans le comté de Limba en 2016. Cet exemple illustre bien comment les résultats spécifiques évalués au regard des critères d'efficacité et d'impact varient selon la définition des objectifs de l'intervention.

Évaluation de l'impact

L'étude avait pour objet d'évaluer spécifiquement l'impact d'un type particulier d'intervention visant à lutter contre la violence à l'égard des femmes et des filles au Libéria, qui consistait à intégrer une composante de transfert en espèces dans une intervention ayant pour but à la fois de contribuer au renforcement économique et de transformer les rapports sociaux entre les femmes et les hommes au bénéfice des très jeunes adolescentes.

L'évaluation a recueilli des informations initiales et finales détaillées sur l'étendue de la violence à l'encontre des jeunes femmes et des filles et sur les effets spécifiques de l'intervention, à la fois pendant l'intervention et un an après. Les résultats ont été mesurés au moyen d'un essai contrôlé randomisé par grappes au niveau des villages. L'étude était conçue de manière à évaluer 1) l'impact global du programme, comparé à un groupe contrefactuel (le groupe de contrôle) et 2) l'efficacité de l'intégration d'un paiement encourageant la participation (« programme GE+ »), en particulier pour mesurer si l'octroi d'incitations financières aux filles présentait des avantages en termes de protection et d'autonomisation, réduisant par là le risque de violence sexuelle, à travers une éventuelle participation accrue des filles aux sessions du programme.

Les effets sur l'autonomisation et l'évolution des normes et des attitudes sociales concernant la place des filles dans la société et la violence sont des exemples d'impacts porteurs de transformation.

Source : Hallman et al. (2018^[22]), *Girl Empower Impact Evaluation: Mentoring and Cash Transfer Intervention to Promote Adolescent Wellbeing in Liberia*, <https://www.rescue.org/sites/default/files/document/4346/girlempowerimpacetevaluation-finalreport.pdf>

Viabilité/durabilité

Définition de la viabilité/durabilité :

Viabilité/durabilité : Les bénéfices seront-ils durables ?

Mesure dans laquelle les bénéfices nets de l'intervention perdureront ou sont susceptibles de perdurer.

Remarque : Sont englobés l'examen des capacités financières, économiques, sociales, environnementales et institutionnelles des systèmes nécessaires à la continuité des bénéfices nets dans le temps. Cela comprend les analyses de la résilience, des risques et des arbitrages potentiels entre priorités. Selon le moment où l'évaluation est effectuée, ce processus peut consister à analyser le flux réel de bénéfices nets ou à estimer la probabilité qu'ils perdurent à moyen et long terme.

Qu'est-ce que la viabilité/durabilité et pourquoi est-elle importante ?

L'évaluation de la viabilité/durabilité permet aux évaluateurs de déterminer si les bénéfices d'une intervention perdureront sur le plan financier, économique, social et environnemental. La viabilité/durabilité pose la question « Les bénéfices se maintiendront-ils dans le temps ? » Si la notion sous-jacente de bénéfices qui perdurent demeure, le critère est à la fois plus concis et plus large dans sa portée que dans sa définition antérieure⁸. La viabilité/durabilité englobe plusieurs éléments d'analyse – financiers, économiques, sociaux et environnementaux – et il convient de prêter attention à l'interaction entre ces éléments.

Il peut y avoir confusion entre la viabilité/durabilité au sens de la continuation des résultats, et la durabilité environnementale ou l'utilisation des ressources pour les générations futures. Même si la durabilité environnementale est une préoccupation (et peut être examinée au titre de plusieurs critères, notamment la pertinence, la cohérence, l'impact et la viabilité/durabilité), elle ne constitue pas en tant que telle l'objet principal des critères ; lorsqu'ils décrivent la viabilité/durabilité, les évaluateurs devraient indiquer clairement la façon dont ils interprètent ce critère.

La viabilité/durabilité devrait être examinée à chaque étape de la chaîne de résultats et du cycle d'une intervention. Les évaluateurs devraient également appréhender cette notion en lien avec celles de résilience et d'adaptation dans des environnements dynamiques et complexes. Cela inclut la viabilité/pérennité des ressources (financières ou autres) une fois l'intervention achevée ainsi que la durabilité/pérennité des impacts dans le contexte plus large de l'intervention. Ainsi, une évaluation pourrait chercher à déterminer si une intervention a pris en compte les capacités des partenaires et favorisé l'appropriation au début de la période de mise en œuvre, et s'il existait une volonté et la capacité de pérenniser le financement à la fin de l'intervention. En règle générale, les évaluateurs peuvent examiner les conditions qui favorisent la viabilité/durabilité au niveau de la conception de l'intervention et de ses activités, et déterminer si des adaptations ont été effectuées, le cas échéant.

Les évaluateurs ne devraient pas considérer la viabilité/durabilité du seul point de vue des donateurs et des apports de financement extérieurs. Les commanditaires devraient également envisager d'évaluer cet aspect avant le début d'une intervention, ou pendant que le financement ou les activités sont en cours. Lorsqu'ils évaluent la viabilité/durabilité, les évaluateurs devraient : 1) tenir compte des bénéfices nets, c'est-à-dire de la valeur globale des bénéfices permanents de l'intervention, y compris les éventuels coûts récurrents, et 2) analyser tout arbitrage potentiel ainsi que la résilience des capacités/systèmes sous-tendant la continuation des bénéfices. Un arbitrage peut être nécessaire, par exemple, entre la viabilité budgétaire des bénéfices et la viabilité politique (c'est-à-dire le maintien du soutien des instances politiques).

L'évaluation de la viabilité/durabilité fournit des informations précieuses sur la continuation (probable) des bénéfices nets d'une intervention sur le moyen à long terme, ce qui se révèle très difficile dans la pratique, comme l'ont montré diverses méta-analyses. À titre d'illustration, dans certains secteurs comme l'eau et l'assainissement ou pour certains types d'intervention comme le développement d'impulsion locale, les bénéfices s'estompent souvent au bout d'un certain temps. Dans ce cas, l'évaluation peut avoir pour rôle d'examiner minutieusement les hypothèses de la théorie du changement relatives à la manière d'assurer la viabilité/durabilité des résultats (Mansuri et Rao, 2013^[23] ; White, Menon et Waddington, 2018^[24]).

Une évaluation qui examine attentivement ces différents aspects pourra livrer d'importants éclairages sur la manière dont les interventions peuvent planifier et mettre en œuvre un changement qui garantisse un développement durable à l'avenir. Les enseignements tirés peuvent mettre en relief la possibilité d'étendre, dans le contexte en question, les mesures de viabilité mises en place dans le cadre de l'intervention, ou de les reproduire dans d'autres contextes.

La planification de la sortie constitue un aspect essentiel de la viabilité/durabilité. Les évaluations devraient déterminer si une stratégie de sortie appropriée a été élaborée et appliquée, ce qui garantirait la continuation des effets positifs, y compris, mais sans s'y limiter, les considérations financières et en matière de capacités. Dans le cas d'une évaluation *ex post*, l'évaluateur peut également examiner si la stratégie de sortie prévue a été dûment mise en œuvre pour garantir la continuation des effets positifs comme anticipé, tout en s'adaptant aux évolutions du contexte, comme le montrent les exemples ci-après.

Les récentes méta-évaluation (Noltze, Euler et Verspohl, 2018^[25]) et synthèse d'évaluation (Noltze, Euler et Verspohl, 2018^[26]) de l'Institut allemand pour l'évaluation du développement (DEval) constituent une ressource utile pour mieux comprendre la viabilité/durabilité et relever des défis communs. Ces deux études mettent en évidence les différents éléments d'analyse de cette notion qui peuvent être examinés dans une évaluation. La méta-évaluation plaide fortement en faveur d'une évaluation de la viabilité/durabilité qui intègre les principes des ODD, et souligne les domaines dans lesquels ce type d'évaluation peut apporter une valeur ajoutée. Il s'agit notamment d'analyser comment l'identification et l'évaluation des effets non intentionnels d'un projet ainsi que des interactions ou des arbitrages entre les différentes dimensions de la viabilité/durabilité peuvent favoriser l'apprentissage et l'exercice de la redevabilité lorsque ce critère est appliqué.

Comprendre la viabilité/durabilité : éléments d'analyse

Pour comprendre la définition de la viabilité/durabilité, il faut en comprendre les composantes que sont la création d'un environnement propice, la continuation des effets positifs, et les risques et les arbitrages.

Créer un environnement propice au développement durable

Les évaluations peuvent examiner la manière dont une intervention a contribué, de multiples façons, à améliorer l'environnement propice au développement, y compris la manière dont elle a permis de renforcer les systèmes, les institutions ou les capacités pour soutenir les activités futures en matière de développement ou d'action humanitaire. Cette approche incite les évaluateurs à prendre en compte les capacités des partenaires au développement qui ont été créées ou renforcées sous l'effet de l'intervention, ainsi que la résilience acquise face aux évolutions et aux chocs extérieurs – garantissant ainsi que les bénéfices nets perdureront (voir plus haut).

Une intervention peut, par exemple, améliorer l'environnement propice au développement à travers les contributions suivantes : renforcement des capacités (au niveau individuel, local ou institutionnel) ; renforcement de l'appropriation ou de la volonté politique ; hausse des engagements financiers ou budgétaires nationaux (et, le cas échéant, infranationaux) ; changement de politique ou de stratégie ; réforme législative ; réformes institutionnelles ; réformes de la gouvernance ; amélioration de la

transparence des dépenses publiques ou amélioration des processus de consultation publique dans la planification du développement.

Continuation des effets positifs : viabilité/durabilité réelle et potentielle

La viabilité/durabilité peut être évaluée sur différentes périodes. Les évaluateurs peuvent évaluer à la fois la viabilité/durabilité réelle (c'est-à-dire la continuation des bénéfices nets découlant de l'intervention qui sont déjà manifestes) et la viabilité/durabilité potentielle (c'est-à-dire les bénéfices nets pour les principales parties prenantes qui sont susceptibles de perdurer). Les évaluateurs devraient réfléchir avec soin aux approches d'évaluation qui conviennent pour évaluer ces deux aspects de la viabilité/durabilité, en fonction du calendrier de l'évaluation et des délais prévus pour obtenir les bénéfices escomptés. Concernant les évolutions de plus vaste portée, beaucoup d'entre elles nécessiteront de longues années, voire des décennies, avant de se réaliser pleinement.

Dans le cadre de l'évaluation de la viabilité/durabilité réelle, l'évaluateur peut examiner dans quelle mesure les effets positifs générés par l'intervention ont manifestement perduré pour les parties prenantes clés, notamment les bénéficiaires ciblés, une fois l'intervention achevée. Les évaluateurs peuvent également s'attacher à déterminer si des possibilités ont été identifiées, anticipées et planifiées pour contribuer à ce que les effets positifs perdurent, et s'il a été fait de même pour d'éventuels obstacles. Ce processus peut étayer les conclusions mettant en évidence la capacité d'une intervention à s'adapter en tant que de besoin.

L'examen de la viabilité/durabilité potentielle suppose une approche légèrement différente. Il s'agit d'évaluer la probabilité que les effets positifs actuels ou prévus de l'intervention perdurent, généralement en partant du principe que les conditions du moment restent les mêmes. L'évaluation devra dans ce cas déterminer le niveau de stabilité et la permanence relative de tout effet positif obtenu, ainsi que les conditions nécessaires à sa pérennité – viabilité/durabilité institutionnelle, économique et financière, environnementale, politique, sociale et culturelle.

Risques et arbitrages potentiels

L'évaluation de la viabilité/durabilité suppose d'examiner non seulement la probabilité que les effets positifs perdurent, mais aussi les risques potentiels et les coûts récurrents associés à une intervention. Par conséquent, les responsables d'évaluation devraient tenir compte des facteurs susceptibles d'améliorer la viabilité/durabilité des bénéfices nets sur la durée, ainsi que des facteurs susceptibles d'y faire obstacle. L'examen des risques liés à la viabilité/durabilité d'une intervention peut nécessiter d'évaluer dans quelle mesure il existe des facteurs contextuels positifs ou négatifs, identifiables ou prévisibles, susceptibles d'influer sur la pérennité des résultats de l'intervention.

D'où la question des arbitrages à opérer, qui est un élément important des critères révisés. L'évaluation des arbitrages associés à une intervention incite à examiner l'arbitrage entre impact immédiat et effets ou coûts potentiels à plus long terme, ainsi que les arbitrages entre les aspects financiers, économiques, sociaux et environnementaux. À titre d'illustration, il peut ressortir d'une évaluation qu'une intervention a soutenu la croissance économique, mais que cette croissance n'est pas durable en raison de ses importants coûts environnementaux, qui peuvent avoir un impact négatif sur la croissance économique à long terme. Un tel constat s'inscrit dans le droit fil de la définition du développement durable consacrée par les ODD et élargit le champ de l'évaluation de sorte que ne soit pas simplement examinée la probabilité d'une continuation des effets positifs d'une intervention.

Liens avec les autres critères

La viabilité/durabilité est étroitement liée aux autres critères.

- La viabilité/durabilité est liée à la pertinence, le niveau de pertinence pour les principales parties prenantes étant un facteur essentiel qui détermine leur appropriation et leur adhésion aux bénéfices ultimes, étapes clés pour assurer la viabilité/durabilité.
- De même, la cohérence peut fournir des indications utiles sur la viabilité/durabilité, car elle examine d'autres interventions menées dans un contexte donné, qui pourraient renforcer, ou compromettre, les bénéfices de l'intervention sur la durée.
- Efficacité et impact : l'évaluation de la continuation des résultats repose d'abord sur le fait que les résultats ont été atteints (efficacité) et ensuite, sur le fait que des effets de plus vaste portée ont été démontrés (impact). Par conséquent, l'efficacité et l'impact peuvent être considérés comme des critères prédominants pour examiner la viabilité/durabilité, car si leur analyse ne montre pas que l'intervention produit des réalisations, des résultats ou des impacts, il n'y aura pas de bénéfices explicites à pérenniser. L'Encadré 4.19 fournit un exemple de la manière dont l'impact et la viabilité/durabilité peuvent être examinés conjointement. Examiner les synergies entre l'impact, l'efficacité et la viabilité/durabilité en évaluant les conditions qui sont suffisantes et nécessaires pour que les résultats perdurent une fois l'intervention achevée permet aux évaluateurs d'analyser l'efficacité et l'impact à plus long terme.
- Les préoccupations en matière d'efficacité peuvent également compromettre la pérennité des bénéfices. La viabilité/durabilité peut être compromise, par exemple, lorsque les coûts à court terme guident la prise de décision, au détriment des effets à plus long terme.

Intégrer la dimension de l'inclusion

La définition révisée de la viabilité/durabilité et la remarque associée attirent l'attention sur les dimensions « financière, économique, environnementale et sociale » de cette notion et sur la manière dont ces dimensions contribuent à ce que les résultats de l'intervention apportent des bénéfices constants et à long terme. Les évaluateurs devraient examiner la manière dont la continuation des bénéfices a été planifiée ou, s'il s'agit d'une évaluation *ex post*, dont elle est manifeste pour les différents groupes de bénéficiaires. À cet égard, l'accent devrait être mis sur le principe consistant à « ne laisser personne de côté » et sur la manière dont les groupes marginalisés tirent des bénéfices positifs permanents, ainsi que sur les arbitrages éventuellement opérés entre différentes catégories de population.

Il est également pertinent que les évaluateurs examinent dans quelle mesure l'intervention a créé un environnement propice à un développement inclusif et équitable, en s'intéressant aux problèmes systémiques sous-jacents (« traiter la maladie, pas seulement les symptômes ») au titre des critères d'impact et de viabilité/durabilité. Les questions d'appropriation et d'autonomisation des femmes sont importantes à cet égard. La viabilité/durabilité des systèmes exige un renforcement des capacités. Les évaluateurs devraient donc être en mesure de déterminer quelles sont les groupes dont les capacités ont été renforcées et quels sont les liens avec les systèmes et les structures inégalitaires en place. Ils devraient se poser la question de savoir si les différentes catégories de parties prenantes possèdent les capacités et affichent l'engagement nécessaires pour créer et préserver à moyen et long terme un environnement propice à l'égalité des sexes et à l'autonomisation des femmes. Si tel n'est pas le cas, ils devraient déterminer quels sont les obstacles qui se dressent sur cette voie et si ceux-ci relèvent du champ d'action de l'intervention.

Difficultés posées par l'évaluation de la viabilité/durabilité et solutions proposées

Le tableau ci-après décrit certaines difficultés rencontrées lors de l'évaluation de la viabilité/durabilité, liées au calendrier, à l'absence d'effets positifs et à d'autres facteurs compromettant la viabilité/durabilité, et propose des solutions à l'intention des évaluateurs comme des responsables d'évaluation.

Tableau 4.6. Difficultés posées par l'évaluation de la viabilité/durabilité

| Difficulté | Solutions : évaluateurs | Solutions : responsables d'évaluation |
|---|---|---|
| L'intervention n'a pas atteint les résultats escomptés (évalués au titre du critère d'efficacité), n'a pas eu d'effets positifs inattendus ou n'a pas contribué à l'impact. | Dans le cas où l'intervention n'a apporté aucun bénéfice, l'analyse de la viabilité/durabilité devient superflue. Les évaluateurs devraient être clairs et ouverts quant aux limites de l'évaluation de la viabilité/durabilité et devraient envisager de se concentrer sur d'autres critères. | Les responsables devraient envisager d'adapter le périmètre de l'évaluation et de réorienter les ressources vers d'autres critères, y compris l'analyse des raisons pour lesquelles les bénéfices n'ont pas été réalisés. |
| Calendrier de l'évaluation. | Les évaluations sont souvent réalisées pendant la durée de vie de l'intervention ou à un moment où la viabilité/durabilité n'est pas encore évidente. Par conséquent, les évaluateurs devraient envisager de se concentrer sur les conditions de la viabilité/durabilité si l'intervention est toujours en cours de mise en œuvre ou vient de s'achever. Ils devraient effectuer une évaluation robuste de la viabilité/durabilité potentielle en recensant les facteurs qui, dans l'environnement opérationnel, pourraient y contribuer. En ce sens, l'analyse sous l'angle de la problématique femmes-hommes est essentielle et fournira des informations utiles et pertinentes sur la cohérence des évolutions structurelles opérées et leur incidence sur la réduction des inégalités. | Si l'évaluation a lieu alors que l'intervention est encore en cours, les responsables devraient préciser dans quelle mesure ils donnent la priorité aux considérations de viabilité/durabilité et quel est l'objectif d'un tel examen. Par exemple, les constatations alimenteront-elles la planification de la sortie/transition ? |
| Évaluation des facteurs susceptibles d'influer sur la viabilité/durabilité. | Les facteurs contextuels qui favorisent (ou compromettent) la viabilité/durabilité potentielle peuvent être évalués de manière qualitative et quantitative. Il s'agit notamment, mais pas exclusivement, de l'appropriation par les parties prenantes et de leur engagement, de la capacité d'absorption, de la volonté politique et de la disponibilité des ressources nationales. | Les responsables peuvent apporter leur concours pour déterminer auprès de quelles parties prenantes pertinentes mener des entretiens ou des enquêtes. |

Exemples d'évaluation de la viabilité/durabilité

Cette partie présente des exemples d'évaluation de la viabilité/durabilité en lien avec un projet de développement rural en Afghanistan, un programme de soutien budgétaire général et un programme relatif à la santé maternelle.

Encadré 4.17. Examen de la viabilité/durabilité dans une évaluation *ex post* d'un projet japonais de développement rural intercommunal en Afghanistan

Le projet de développement rural intercommunal a été mis en œuvre dans le cadre de l'initiative Ogata de l'Agence japonaise de coopération internationale (JICA). Il visait à établir un modèle de développement rural piloté au niveau local, susceptible de contribuer à la reconstruction et au développement rural dans les provinces de Balkh, Bamyan et Kandahar en Afghanistan. L'évaluation *ex post* a examiné la viabilité/durabilité de cette intervention en lien avec les efforts de reconstruction consécutifs à la guerre civile en Afghanistan.

Après avoir déterminé que l'intervention avait été mise en œuvre avec succès et avait atteint ses objectifs, l'évaluation a examiné la pérennité des bénéfices obtenus. Dans le cadre de l'analyse s'est posée la question de savoir si les conditions politiques, institutionnelles, techniques et financières étaient en place pour permettre aux effets positifs de perdurer. Parmi ces conditions figuraient notamment la capacité et la volonté du gouvernement afghan de poursuivre la mise en œuvre après l'achèvement du projet.

L'évaluation a conclu que l'intervention avait mis en place des systèmes de suivi en coopération avec d'autres donateurs, ainsi qu'une base budgétaire reposant sur les apports des donateurs, ce qui

complétait et renforçait les capacités du système national du pays touché par le conflit. Elle a également confirmé que le gouvernement afghan était déterminé à poursuivre le développement rural piloté au niveau local en regroupant les conseils locaux de développement, avec le soutien des donateurs. Le projet a également été intégré au Programme de solidarité nationale. Par conséquent, l'évaluation a conclu que les conditions politiques, institutionnelles, techniques et financières nécessaires à la continuation des effets positifs étaient en place et qu'il n'existait aucun risque apparent.

Source : Watanabe (2016^[27]), *Ex-Post Evaluation of Technical Co-operation Project "Inter-Communal Rural Development Project"*, https://www2.jica.go.jp/en/evaluation/pdf/2015_0603847_4.pdf

Encadré 4.18. Leçons sur la viabilité/durabilité tirées de l'évaluation du retrait du programme de soutien budgétaire général au Malawi, en Ouganda, au Rwanda et en Zambie

Cette étude, qui constitue un exemple intéressant d'utilisation réfléchie du critère de viabilité/durabilité dans une évaluation, examinait dans quelle mesure l'intervention avait dûment posé les jalons nécessaires pour que ses effets positifs perdurent après le retrait du programme de soutien budgétaire, et si elle contribuait à la création d'un environnement propice au développement. Elle été réalisée à la suite du retrait du soutien budgétaire général opéré à partir de 2010 environ en raison d'une série de scandales dans les pays bénéficiaires et de pressions dans les pays donateurs pour que la reddition de comptes soit assurée. Ce revers faisait suite au mouvement en faveur de mesures de soutien budgétaire, apparu au début des années 2000 en réaction à la Déclaration de Paris sur l'efficacité de l'aide (2005). Le soutien budgétaire était largement considéré comme un moyen efficace et durable de contribuer aux résultats en matière de développement. Plusieurs évaluations et études de synthèse montraient, données à l'appui, que le soutien budgétaire avait eu un effet positif sur des résultats importants en matière de développement, comme l'augmentation des dépenses en faveur des pauvres, l'accroissement de la scolarisation et l'amélioration de la gestion des finances publiques.

L'évaluation a pris la forme d'une étude de cas comparative au Malawi, en Ouganda, au Rwanda et en Zambie et visait à déterminer si des stratégies appropriées avaient été mises en œuvre pour garantir la continuation des effets positifs associés au soutien budgétaire général après son retrait. Elle a mis en lumière le fait que, même si le niveau global d'aide était resté constant dans la plupart des pays à la suite de la décision de supprimer le soutien budgétaire général, les structures établies avaient pour la plupart disparu. La majorité des réalisations positives avaient été négativement affectées, voire réduites à néant. Le financement indépendant de type projet s'est révélé la nouvelle modalité dominante dans les portefeuilles d'aide bilatérale, le financement conjoint étant limité à quelques secteurs et programmes. Cette situation a entraîné des externalités négatives et une forte fragmentation de l'aide. Partant, il a été difficile de favoriser une appropriation nationale ou de s'attaquer à des problèmes systémiques ou plus généraux en matière de gouvernance.

Source : Orth, Birsan et Gotz (2018^[28]), *The Future of Integrated Policy-Based Development Cooperation: Lessons from the Exit from General Budget Support in Malawi, Rwanda, Uganda and Zambia*, http://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/DEval_EN_The%20Future%20of%20Integrated%20Policy-Based%20Development%20Cooperation..pdf

Encadré 4.19. Évaluer l'impact et la viabilité/durabilité du programme du DFID en faveur de la santé maternelle

Menée par l'Independent Committee on Aid Impact (ICAI, commission indépendante sur l'impact de l'aide), cette évaluation des résultats obtenus par le Department for International Development (DFID) du Royaume-Uni dans le domaine de l'amélioration de la santé maternelle dans différents pays illustre comment l'impact et la viabilité/durabilité peuvent être considérés simultanément. Les examens de l'ICAI suivent une approche spécifique qui, dans le cas présent, accordait la priorité à trois grands domaines, à savoir l'impact, l'équité et la viabilité/durabilité. Les questions correspondantes étaient notamment les suivantes :

- Impact : dans quelle mesure le DFID a-t-il maximisé l'impact de ses programmes relatifs à la santé maternelle ? Quelle est la solidité de ses affirmations concernant les résultats obtenus en matière de santé maternelle aux niveaux de la planète, des pays et des programmes ? Dans quelle mesure ses investissements tiennent-ils compte des données mondiales sur les éléments qui fonctionnent pour améliorer la santé maternelle, et dans quelle mesure ses programmes ont-ils produit les résultats requis pour améliorer la santé maternelle ?
- Équité : avec quelle efficacité les programmes ont-ils ciblé les femmes difficiles à atteindre et marginalisées ? Dans quelle mesure le DFID contribue-t-il au renforcement des systèmes de santé et à l'amélioration de l'accès des femmes à ces systèmes ?
- Viabilité/durabilité : les résultats du DFID en matière de santé maternelle sont-ils susceptibles de se révéler durables ? Dans quelle mesure le DFID a-t-il soutenu la mise en place et la mise en œuvre de politiques et d'institutions nationales pour garantir des résultats durables ?

Dans le présent contexte, l'équité englobe la prise en compte directe des effets spécifiques sur les femmes pauvres et jeunes, conformément à la nature de l'intervention. Ce type de questions, portant sur les résultats différenciés et sur des groupes de population particuliers, peuvent être abordées sous différents angles dans le cadre des nouvelles définitions de la pertinence, de l'efficacité et de l'impact. La viabilité/durabilité est considérée en termes de « probabilité » de pérennité des résultats, étant donné que la durabilité réelle (*ex post*) ne peut être observée que sur une période plus longue (dans le futur).

Dans ce cas, la méthodologie consiste à réaliser une évaluation indépendante (dans le respect du rôle et du mandat de l'ICAI) des affirmations du DFID concernant les résultats de ses activités d'amélioration de la santé maternelle, y compris ses travaux de modélisation. Elle s'appuie sur des documents, des entretiens et des visites sur le terrain.

Source : ICAI (2018^[29]), *Assessing DFID's results in improving Maternal Health: An impact review*, <https://icai.independent.gov.uk/wp-content/uploads/ICAI-review-Assessing-DFIDs-results-in-improving-Maternal-Health-.pdf>

Références

- ADE (2019), *Joint Evaluation of the Integrated Solutions Model in and around Kalobeyei, Turkana, Kenya*, UNHCR and Danida, https://um.dk/en/danida-en/results/eval/eval_reports/publicationdisplaypage/?publicationid=dd54bef1-1152-468c-974b-c171fbc2452d (consulté le 11 janvier 2021). [36]
- Affaires mondiales Canada (2019), *Évaluation de l'aide à la reconstruction en cas de catastrophe naturelle aux Philippines 2013-14 to 2018-19*, Affaires mondiales Canada, <https://www.international.gc.ca/gac-amc/publications/evaluation/2019/endra-earcn-philippines.aspx?lang=fra> (consulté le 12 janvier 2021). [8]
- Arghiros, D. et al. (2017), *Making it Count: Lessons from Australian Electoral Assistance 2006-16*, Gouvernement de l'Australie, Département des affaires étrangères et du commerce, <http://www.oecd.org/derec/australia/australia-electoral-assistance-2006-2016.pdf> (consulté le 11 janvier 2021). [11]
- Baltzer, K. et H. Hansen (2011), *Agricultural Input Subsidies in Sub-Saharan Africa: Evaluation Study*, DANIDA, Agence danoise de développement international, ministère des affaires étrangères du Danemark,, <http://www.oecd.org/derec/49231998.pdf> (consulté le 11 janvier 2021). [15]
- Bamberger, M., J. Vaessen et E. Raimondo (2015), *Dealing With Complexity in Development Evaluation - A Practical Approach*, https://www.betterevaluation.org/en/resources/dealing_with_complexity_in_development_evaluation (consulté le 11 janvier 2021). [41]
- Beck, T. (2016), *Evaluating Humanitarian Action using the OECD-DAC Criteria*, ALNAP, <https://www.alnap.org/help-library/evaluating-humanitarian-action-using-the-oecd-dac-criteria> (consulté le 11 janvier 2021). [32]
- Befani, B. et J. Mayne (2014), « Process Tracing and Contribution Analysis: A Combined Approach to Generative Causal Inference for Impact Evaluation », *IDS Bulletin*, vol. 45/6, pp. 17-36, <https://doi.org/10.1111/1759-5436.12110>. [48]
- Belcher, B. et M. Palenberg (2018), « Outcomes and Impacts of Development Interventions », *American Journal of Evaluation*, vol. 39/4, pp. 478-495, <https://doi.org/10.1177/1098214018765698>. [47]
- BMZ and DEval (2020), *Evaluation Criteria for German Bilateral Development Co-operation*. [40]
- Bryld, E. (2019), *Evaluation of Sida's Support to Peacebuilding in Conflict and Post-Conflict Contexts: Somalia Country Report*, Asdi, https://publikationer.sida.se/contentassets/1396a7eb4f934e6b88e491e665cf57c1/eva2019_5_62214en.pdf (consulté le 11 janvier 2021). [19]
- Chambers, R. et al. (2009), « Méthodologie de l'évaluation d'impact : présentation de différentes approches », n° 4, 3ie, https://www.3ieimpact.org/sites/default/files/2019-05/3ie_working_paper_4_french_0.pdf (consulté le 11 janvier 2021). [46]
- D'Errico, S., T. Geoghe et I. Piergallini (2020), *Evaluation to connect national priorities with the SDGs | Publications Library*, IIED, <https://pubs.iied.org/17739IIED> (consulté le 22 février 2021). [38]

- Danida (2019), *Evaluation of Water, Sanitation and Environment Programmes in Uganda (1990-2017)*, Evaluation Department, Ministry of Foreign Affairs of Denmark, <http://www.oecd.org/derec/denmark/denmark-1990-2017-wash-environment-uganda.pdf> (consulté le 11 janvier 2021). [35]
- Davis, R. (2013), « Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations », n° 40, DIFD, <https://www.gov.uk/government/publications/planning-evaluability-assessments> (consulté le 12 janvier 2021). [42]
- Economía Urbana et IPSOS (2019), *Evaluación de operaciones del programa « más familias en acción » y de resultados del componente de bienestar comunitario*, [Evaluation of the "More Families in Action" programme], Departamento Nacional de Planeación, Bogotá D.C., https://colaboracion.dnp.gov.co/CDT/Sinergia/Documentos/Evaluacion_MFA_Informe_Resultados.pdf (consulté le 15 janvier 2021). [18]
- Eurecna Spa (2020), *Bolivia - Evaluation of Health Initiatives (2009-2020)*, Ministère des Affaires étrangères et de la Coopération internationale de l'Italie, http://www.oecd.org/derec/italy/evaluation-report-of-health-initiatives-in-Bolivia-2009_2020.pdf (consulté le 11 janvier 2021). [5]
- FAO (2020), *Evaluation of « Improving farmer livelihoods in the dry zone through improved livestock health, productivity and marketing »*, Organisation des Nations Unies pour l'alimentation et l'agriculture, Rome, <http://www.fao.org/3/ca8463en/ca8463en.pdf> (consulté le 11 janvier 2021). [4]
- Gertler, P. et al. (2016), *Impact Evaluation in Practice: Second Edition*, Groupe de la Banque mondiale, Washington, D.C., <https://publications.iadb.org/en/impact-evaluation-practice-second-edition> (consulté le 12 janvier 2021). [45]
- GNUE (2013), *Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management*, <http://www.uneval.org/document/detail/1433> (consulté le 12 janvier 2021). [44]
- Hallman, K. et al. (2018), *Girl Empower Impact Evaluation: Mentoring and Cash Transfer Intervention to Promote Adolescent Wellbeing in Liberia*, International Rescue Committee, <https://www.rescue.org/sites/default/files/document/4346/girlempowerimpacetevaluation-finalreport.pdf> (consulté le 12 janvier 2021). [22]
- Helle, E. et al. (2018), *From Donors to Partners? Evaluation of Norwegian Support to Strengthen Civil Society in Developing Countries through Norwegian Civil Society Organisations*, Norad Norwegian Agency for Development Cooperation, https://www.norad.no/globalassets/filer-2017/evaluating/1.18-from-donor-to-partners/1.18-from-donors-to-partners_main-report.pdf (consulté le 11 janvier 2021). [34]
- ICAI (2018), *Assessing DFID's Results in Improving Maternal Health: An Impact Review*, The Independent Commission for Aid Impact, <https://icai.independent.gov.uk/wp-content/uploads/ICAI-review-Assessing-DFIDs-results-in-improving-Maternal-Health-.pdf> (consulté le 11 janvier 2021). [29]
- IDD and Associates (2006), *A Joint Evaluation of General Budget Support Evaluation of General Budget Support: Synthesis Report*, <https://www.oecd.org/development/evaluation/dcdndep/37426676.pdf> (consulté le 11 janvier 2021). [2]

- IDEV (2018), *Cabo Verde : Évaluation de la stratégie et du programme pays de la Banque 2008–2017 : Résumé analytique*, Évaluation indépendante du développement, Banque africaine de développement, <https://idev.afdb.org/sites/default/files/Evaluations/2020-03/Executive%20Summary%20Cabo%20Verde%20CSPE%20%28Fr%29.pdf> (consulté le 11 janvier 2021). [12]
- IEG (2008), *The Welfare Impact of Rural Electrification: A Reassessment of the Costs and Benefits*, Banque mondiale, <https://doi.org/10.1596/978-0-8213-7367-5>. [14]
- IOB (2017), *Tackling major water challenges: Policy review of Dutch development aid policy for improved water management*, Policy and Operations Evaluation Department, Ministry of Foreign Affairs, Netherlands, <https://english.iob-evaluatie.nl/publications/policy-review/2017/12/01/418-%E2%80%93-iob-%E2%80%93-policy-review-of-dutch-development-aid-policy-for-improved-water-management-2006-2016-%E2%80%93-tackling-major-water-challenges> (consulté le 12 janvier 2021). [16]
- Leeuw, F. et J. Vaessen (2009), *Impact Evaluations and Development: NoNIE Guidance on Impact Evaluation*, The Networks on Impact Evaluation, Washington D.C., <http://search.oecd.org/dac/evaluation/dcdndep/47466906.pdf> (consulté le 12 janvier 2021). [43]
- Leppert, G. et al. (2018), *Impact, Diffusion and Scaling-Up of a Comprehensive Land-Use Planning Approach in the Philippines: From Development Cooperation to National Policies*, German Institute for Development Evaluation (DEval), Bonn, https://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/Zusammenfassung_D_utsch_%20DEval-2018_Philippinen_final_web-2.pdf (consulté le 12 janvier 2021). [21]
- Mansuri, G. et V. Rao (2013), *Localizing Development*, Banque mondiale, <https://doi.org/10.1596/978-0-8213-8256-1>. [23]
- Ministère des affaires étrangères de la Slovénie (2017), *Evaluation of Slovenia's Development Cooperation with Montenegro 2013-2016: Final Report*, Slovenia's Development Cooperation, <https://www.gov.si/assets/ministrstva/MZZ/Dokumenti/multilaterala/razvojno-sodelovanje/Development-cooperation-with-Montenegro-evaluation-final-report.pdf> (consulté le 12 janvier 2021). [9]
- Ministère des affaires étrangères du Japon (2019), *Japan ODA Evaluation Guidelines*, https://www.mofa.go.jp/policy/oda/evaluation/basic_documents/pdfs/guidelines11th.pdf (consulté le 18 février 2021). [6]
- Molund, S. et G. Schill (2004), *Looking Back, Moving Forward Sida Evaluation Manual*, Sida, <https://www.oecd.org/derec/sweden/35141712.pdf> (consulté le 11 janvier 2021). [17]
- Noltze, M., M. Euler et I. Verspohl (2018), *Evaluation Synthesis of Sustainability in German Development Cooperation*, German Institute for Development Evaluation (DEval), Bonn, http://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/DEval_Evaluierungssynthese_EN_web.pdf (consulté le 12 janvier 2021). [26]
- Noltze, M., M. Euler et I. Verspohl (2018), *Meta-Evaluation of Sustainability in German Development Cooperation*, German Institute for Development Evaluation (DEval), Bonn, http://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/DEval_NH_Meta-Evaluierung_EN_web.pdf (consulté le 12 janvier 2021). [25]

- Norad (2018), *Evaluation of Norwegian Efforts to Ensure Policy Coherence for Development*, Norad, <https://www.norad.no/contentassets/4ac3de36fbd4229811a423f4b00acf7/8.18-evaluation-of-norwegian-efforts-to-ensure-policy-coherence-for-development.pdf> (consulté le 11 janvier 2021). [7]
- OCDE (2011), *Approche méthodologique pour les évaluations de l'aide budgétaire*, Réseau du CAD sur l'évaluation du développement, Éditions OCDE, Paris, https://www.oecd.org/dac/evaluation/Approche%20methodologique_aide%20budgetaireFINA L.pdf (consulté le 12 janvier 2021). [3]
- OCDE (2010), *Normes de qualité pour l'évaluation du développement*, Lignes directrices et ouvrages de référence du CAD, Éditions OCDE, Paris, <https://doi.org/10.1787/9789264083912-fr>. [1]
- OCDE (2002), *L'évaluation et l'efficacité de l'aide n° 6 - Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats (en anglais, français et espagnol)*, Éditions OCDE, Paris, <https://doi.org/10.1787/9789264034921-en-fr>. [10]
- OECD (2019), *Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use*, DAC Network on Development Evaluation, OECD Publishing, Paris, <https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf> (consulté le 11 janvier 2021). [31]
- OECD (1991), *Principles for Evaluation of Development Assistance*, Development Assistance Committee, OECD Publishing, Paris, <https://www.oecd.org/development/evaluation/2755284.pdf> (consulté le 11 janvier 2021). [30]
- Ofir, Z. et al. (2016), *Briefing: Five considerations for national evaluation agendas informed by the SDGs*, IIED, London, <https://doi.org/10.3138/cjpe.30.3.02./11>. [39]
- Orth, M., M. Birsan et G. Gotz (2018), *The Future of Integrated Policy-Based Development Cooperation: Lessons from the Exit from General Budget Support in Malawi, Rwanda, Uganda and Zambia*, German Institute for Development Evaluation (DEval), Bonn, http://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/DEval_EN_The%20Future%20of%20Integrated%20Policy-Based%20Development%20Cooperation.pdf (consulté le 12 janvier 2021). [28]
- Palenberg, M. (2011), *BMZ: Tools and Methods for Evaluating the Efficiency of Development Interventions - Managing for Sustainable Development Impact*, BMZ Evaluation Division – German Federal Ministry for Economic Cooperation and Development, <http://www.managingforimpact.org/resource/bmz-tools-and-methods-evaluating-efficiency-development-interventions> (consulté le 12 janvier 2021). [13]
- PEM Consult (2020), *Evaluation of the Danish Strategic Sector Cooperation*, Evaluation, Learning and Quality Department, Danida (Ministère des Affaires étrangères), Danemark, https://um.dk/en/danida-en/results/eval/eval_reports/publicationdisplaypage/?publicationID=CBE77158-1D4D-46E3-A81F-9B918218FAFF (consulté le 12 janvier 2021). [20]
- Schwedersky, T., L. Ahrens et H. Steckhan (2017), *Evaluation of the BMZ Action Plan for the Inclusion of Persons with Disabilities*, DEval, <https://d-nb.info/1186644206/34> (consulté le 11 janvier 2021). [33]

- Watanabe, K. (2016), *Ex-Post Evaluation of Technical Cooperation Project « Inter-Communal Rural Development Project »*, JICA, [27]
https://www2.jica.go.jp/en/evaluation/pdf/2015_0603847_4.pdf (consulté le 12 janvier 2021).
- WFP (2016), *Technical Note: Evaluation Methodology*, DEQAS, World Food Programme, [37]
<https://docs.wfp.org/api/documents/704ec01f137d43378a445c7e52dcf324/download/>
 (consulté le 11 janvier 2021).
- White, H., R. Menon et H. Waddington (2018), « Community-driven development: does it build social cohesion or infrastructure? A mixed-method evidence synthesis », n° 30, 3ie, [24]
<https://www.3ieimpact.org/evidence-hub/publications/working-papers/community-driven-development-does-it-build-social-cohesion> (consulté le 12 janvier 2021).

Notes

¹ C'est-à-dire « les individus, les groupes ou les organisations qui bénéficient de l'intervention de développement, directement ou non, intentionnellement ou non. » D'autres termes peuvent être employés, comme « détenteurs de droits » ou « personnes affectées ».

² Il peut arriver que deux institutions soient engagées dans une même intervention, l'une en qualité de partenaire d'exécution et l'autre en qualité de bailleur de fonds. Il est alors possible d'évaluer la cohérence interne du point de vue du bailleur de fonds mais aussi de celui du partenaire d'exécution.

³ Il convient toutefois de signaler que l'évaluation de l'impact a beaucoup plus de chances d'être utile si elle s'accompagne d'une analyse de la manière dont l'impact est obtenu et des moyens par lesquels il pourrait être renforcé que s'il s'agit d'un simple exercice de redevabilité.

⁴ Un examen détaillé et complet des différentes méthodologies utilisables pour définir et évaluer l'efficacité est proposé par Palenberg (2011[24]) dans un document de travail du BMZ. L'auteur identifie trois niveaux différents : le niveau 0 (décrire certains aspects de l'intervention d'aide relatifs à l'efficacité et émettre un avis en la matière) ; le niveau 1 (déterminer le potentiel d'amélioration de l'efficacité d'une intervention d'aide), qui fournit une image partielle des processus de mise en œuvre, du coût des intrants, de la conversion des intrants en extrants ou de la conversion des extrants en réalisations ; le niveau 2 (évaluer l'efficacité d'une intervention d'aide de manière à pouvoir la comparer avec d'autres interventions ou avec des points de référence). Il s'agit d'une approche globale qui inclut une estimation fiable des principaux coûts et avantages.

En pratique, l'évaluation de niveau 2 est rarement pratiquée. Même la Banque mondiale, pourtant dotée de moyens colossaux, relevait dans un rapport de 2010 que l'analyse coût-avantages était nettement moins pratiquée que par le passé au stade de l'évaluation, même dans les secteurs où elle s'appliquait particulièrement bien. Dans ce rapport, elle insistait sur l'avantage, en termes de rigueur et de liens avec les résultats ultérieurs du projet, d'une analyse aussi approfondie, tout en soulignant les difficultés qu'elle comporte (IEG, 2010 [25]). Le cas d'étude sur l'électrification en milieu rural (IEG, 2008[26]) est un exemple relativement rare d'analyse complète de niveau 2. Les évaluateurs peuvent aussi opter pour des méthodes telles que la prise de décision/modélisation à critères multiples. Il importe également de noter que ces

types d'analyse sont davantage susceptibles d'être appliqués dans certains secteurs – infrastructures, santé et agriculture, par exemple. L'Encadré 4.10 cite un exemple intéressant, en l'occurrence celui des subventions aux intrants dans le secteur agricole, qui se fonde sur une étude de cas nationale et s'appuie sur diverses estimations économiques et données d'enquête.

⁵ La littérature sur ce sujet est abondante, mais on pourra consulter, par exemple : GNUE (2013^[44]) ; Chambers et al. (2009^[46]) ; Leeuw et Vaessen (2009^[43]) ; Belcher et Palenberg (2018^[47]) ; et Gertler et al. (2016^[45])

⁶ Cette notion ne doit pas être confondue avec l'importance (ou significativité) statistique, qui est souvent évoquée dans le cadre de certains types d'évaluations d'impact ; voir Gertler et al. (2016^[45]) sur les calculs de puissance et les concepts techniques correspondants dans les évaluations d'impact quantitatives.

⁷ Voir Befani et Mayne (2014^[48]).

⁸ Dans le *Glossaire* de 2002 (OCDE, 2002^[10]), la viabilité (pérennité, durabilité) est définie comme suit : « Continuation des bénéfices résultant d'une action de développement après la fin de l'intervention. Probabilité d'obtenir des bénéfices sur le long terme. Situation dans laquelle les avantages nets sont susceptibles de résister aux risques. »

Une application réfléchie des critères d'évaluation

La pertinence, la cohérence, l'efficacité, l'efficience, l'impact et la durabilité sont des critères d'évaluation largement utilisés, en particulier dans la coopération internationale pour le développement. Ils permettent de déterminer le mérite ou la valeur de diverses interventions, telles que les stratégies, les politiques, les programmes ou les projets. Ce guide a pour but d'aider les évaluateurs et d'autres à mieux comprendre ces critères et mieux les utiliser. Il commence par décrire ce qu'ils sont et comment ils sont censés être utilisés. Les définitions et les concepts qui sous-tendent chaque critère sont ensuite expliqués. Enfin, des exemples donnent au lecteur des idées concrètes pour leur utilisation.

Les critères ont été définis au début des années 2000 par le Réseau sur l'évaluation du développement (EvalNet) du Comité d'aide au développement (CAD) de l'OCDE. Bien qu'ils aient été largement utilisés dans le domaine de l'évaluation et au-delà, ce document est le premier à en explorer les concepts en détail, à expliquer leur utilisation prévue et à offrir des conseils pratiques. Il reflète l'état actuel de la réflexion et les meilleures pratiques en matière d'évaluation, en s'appuyant sur les contributions d'experts en évaluation de renommée internationale venus d'EvalNet et d'ailleurs.



IMPRIMÉ ISBN 978-92-64-52037-0
PDF ISBN 978-92-64-32295-0



9 789264 520370