

Renforcer la pertinence, l'utilité et la fiabilité des évaluations d'impact à l'AFD (Note de cadrage)

RESUME EXECUTIF

À ce jour, l'AFD a conclu six évaluations d'impact, sur les vingt lancées depuis le début des années 2000. Les connaissances issues de ces recherches évaluatives ont significativement alimenté les débats internationaux sur le développement, en particulier dans les domaines de la microfinance, de la protection sociale et du financement de la santé. Ces expériences ont en outre fait progresser l'ensemble des méthodes d'évaluation de l'agence. Trois études d'impact sont en cours actuellement : sur l'aménagement urbain à Djibouti, la mise à niveau des entreprises au Sénégal et l'accès à l'eau à Uvira (RDC). Trois autres sont en phase d'instruction : sur l'accès à l'eau à Kinshasa, sur la santé en Guinée et sur la santé maternelle et infantile en Mauritanie. Ces dernières sont toutefois compromises, souvent en raison de défaillances des contreparties locales, qui ne disposent pas des compétences et des procédures adaptées pour porter de telles activités de recherche évaluative.

Compte tenu de ces difficultés et de la part importante d'évaluations d'impact avortées – commune à l'ensemble des bailleurs –, la Direction générale et le Comité des évaluations ont mandaté la division de l'évaluation (EVA) pour qu'elle coordonne une démarche permettant d'améliorer le taux de réussite, la pertinence et l'utilité de ces exercices. Il s'agit *in fine* de fournir aux opérationnels des éléments leur permettant d'orienter leurs stratégies et leurs pratiques afin d'améliorer la capacité de l'AFD à atteindre ses objectifs de développement. Il ressort de cette réflexion que, pour être en mesure de mener ces évaluations dans des conditions satisfaisantes, il est nécessaire qu'elles soient engagées dès l'instruction des projets, au travers d'une procédure claire et sur la base de critères partagés entre la direction des opérations (DOE) et EVA. Il est aussi indispensable que la maîtrise d'ouvrage de l'évaluation soit déléguée à EVA si la contrepartie locale en charge du projet ne dispose pas des compétences pour diriger cette étude, et que les procédures de passation de marché s'adaptent aux spécificités de ce type d'exercice, permettant y compris le recours au gré à gré, dans le cadre de partenariats de recherche avec des laboratoires publics ou des instituts nationaux de statistique.

Toutefois, les moyens que l'AFD consacre actuellement aux évaluations d'impact sont insuffisants pour en commanditer directement et, lorsqu'elles sont financées sur budgets des projets, il s'avère souvent impossible de procéder à des délégations de maîtrise d'ouvrage ou à l'établissement de partenariats de recherches. Malgré ces difficultés, la Direction générale de l'AFD a réaffirmé son engagement à mener des évaluations d'impact. Par conséquent, le Comité des évaluations du 9 avril 2014 recommande de fixer un objectif quantitatif modeste, mais atteignable, consistant pour l'AFD à commanditer sur son budget propre approximativement une évaluation tous les deux ans. Sous réserve d'y consacrer le budget requis, l'Agence serait en mesure d'en assurer le pilotage, permettant notamment d'entretenir son expertise interne sur le sujet. Ce rôle doit toutefois être complété par d'autres activités connexes : une fonction de « courtier en connaissances » sur l'impact au profit des opérationnels (revues de littératures, méta-évaluations, animation scientifique...), une médiation auprès des partenaires scientifiques de l'AFD (IRD, CEPREMAP, PES, etc.), afin de les encourager à mener ce type d'études, et une contribution aux débats internationaux sur les méthodes et les approches à privilégier en matière d'évaluation d'impact des interventions de développement.

INTRODUCTION

Dans le cadre de la nouvelle politique d'évaluation de l'AFD et suite aux débats suscités lors de la conférence EUDN de 2012 et aux questionnements de la DG Trésor sur les évaluations d'impact, EVA a animé des réflexions visant à renforcer leur portée, ainsi que leur rapport coût-utilité. En octobre 2013, EVA a présenté au Comité des évaluations un bilan approfondi de cette activité, qui a ensuite été publié dans la collection *À Savoir*¹. Celui-ci proposait deux alternatives : maintenir une activité d'évaluation d'impact au sein de l'AFD ou l'externaliser, par exemple au travers du réseau 3IE (international initiative for impact evaluation). Le Comité des évaluations et la Direction générale se sont prononcés en faveur du maintien de cette activité au sein de l'AFD.

Mettre en œuvre cette directive de manière efficace implique d'adapter la stratégie et les processus organisationnels afférents : critères et séquence d'identification des projets devant faire l'objet d'une évaluation d'impact, financement de ces études, maîtrise d'ouvrage et gouvernance, mode de passation de marchés pour la maîtrise d'œuvre. Cette approche gagnera à s'accompagner d'activités complémentaires : gestion des connaissances au profit des opérationnels, animation de partenariats scientifique et présence dans les débats internationaux sur les méthodes de recherche et d'évaluation.

A. ETAT DES LIEUX DES EVALUATIONS D'IMPACT A L'AFD

1. Définition et méthodes

En matière de développement, la notion d'impact est constamment mobilisée, mais avec des acceptions distinctes, mouvantes et parfois contradictoires. Nous nous référons dans ce document aux évaluations visant, de manière rigoureuse, à identifier, mesurer et comprendre les effets *strictement attribuables* à une intervention au moyen d'un contrefactuel, c'est-à-dire en comparant l'évolution d'une population traitée par rapport à la situation qui aurait prévalu en l'absence d'intervention.

Cette définition, proposée par le Comité des évaluations, diffère de celle fixée en 1991 par le Comité d'aide au développement (CAD) de l'OCDE, qui englobe l'ensemble des « effets à long terme, positifs et négatifs, primaires et secondaires, induits par une action de développement, directement ou non, intentionnellement ou non »². La notion d'impact mobilisée par le CAD est plus large et n'attache pas le même soin à la démonstration de *l'attribution*. Nous ne traitons pas non plus ici des indicateurs agréables fixés par l'AFD, dont certains sont qualifiés d'indicateurs d'impact (par exemple l'amélioration du taux d'alphabétisation ou la diminution du taux de mortalité)³. En effet, ceux-ci traduisent l'évolution de variables que l'intervention vise à améliorer, mais ils ne permettent pas de déterminer dans quelle mesure elle y contribue effectivement. Dans ce document, nous ne nous référons donc qu'à la première définition, soulignée ci-dessus.

Plusieurs méthodes existent pour réaliser une évaluation d'impact :

- les méthodes quantitatives *expérimentales*, qui impliquent qu'avant le projet, le groupe de bénéficiaires et le groupe de contrôle auxquels on les compare soient désignés de façon aléatoire. Plus robustes en théorie, elles s'avèrent souvent difficiles à mettre en œuvre.
- les méthodes quantitatives *quasi-expérimentales* consistent à identifier, sans influencer le mode de sélection du projet, des populations qui semblent comparables aux bénéficiaires. Leur mise en œuvre est donc plus souple et elles sont susceptibles de fournir des résultats robustes. Un « biais de sélection » peut toutefois subsister, c'est-à-dire que des différences structurelles subsistent entre les bénéficiaires et le groupe de contrôle, ce qui fausse la comparaison.
- les méthodes *mixtes* combinent l'une des méthodes ci-dessus avec des études qualitatives, anthropologiques ou socio-économiques. Elles recourent à des techniques d'entretien, d'observation et d'analyse pour trianguler les informations et comprendre les phénomènes mesurés. Cette articulation permet d'améliorer la fiabilité des résultats et d'en tirer des enseignements généralisables.

2. L'action de l'AFD en matière d'évaluation d'impact

Une décennie de recherche évaluative

L'AFD mène des évaluations d'impact depuis les années 2000. Deux premières études ont porté sur l'hévéaculture et la microfinance. En 2006, la division de l'évaluation a été intégrée au département de la recherche, pour mettre cette activité au service de la production de connaissances. Les évaluations d'impact ont alors pris de l'importance au sein de la division, avec le souci de renforcer leur validité scientifique et leur utilité pour la décision stratégique⁴. Des moyens humains et financiers significatifs leur ont été dédiés et deux études particulièrement ambitieuses ont été lancées, privilégiant des approches jugées particulièrement innovantes et rigoureuses, et mobilisant des équipes de recherche prestigieuses.

En 2006 l'agence a commandité au J-PAL Europe (MIT-PSE-CREST) une vaste étude sur le microcrédit au Maroc. Ce type de prêts était alors fréquemment mis en avant comme un outil privilégié pour lutter contre la pauvreté, favoriser l'éducation et la santé, et promouvoir l'émancipation des femmes⁵. Cette efficacité n'avait cependant pas été démontrée, car la diffusion de la microfinance et le fait qu'elle repose sur un enrôlement volontaire constituaient deux obstacles majeurs pour constituer un contrefactuel incontestable. Les multiples recherches menées sur ce sujet faisaient l'objet d'interminables controverses. Evaluer l'impact de l'extension du microcrédit dans des zones rurales marocaines – encore singulièrement peu desservies – en employant une méthode novatrice fondée sur l'expérimentation par assignation aléatoire, semblait *a priori* constituer une opportunité unique pour surmonter les biais potentiels de ce type d'études et apporter une contribution décisive aux connaissances dans ce domaine.

Parallèlement, l'AFD a engagé avec le SEGA (Berkeley) une étude sur l'assurance santé au Cambodge. La protection des risques pose plusieurs défis majeurs en matière d'évaluation. Il est d'une part très difficile de quantifier les risques, en raison de leur caractère intrinsèquement hypothétique. D'autre part, la couverture de ces risques peut générer des comportements de sélection adverse ou de dés-incidentation qui brouillent ses effets positifs. En l'absence de méthode crédible pour en démontrer les effets vertueux, la protection sociale est généralement perçue avant tout comme un coût, dont les bénéfices demeurent évanescents. Là encore, les méthodes expérimentales paraissaient en mesure d'apporter enfin des éléments empiriques tangibles sur des modes de gestion des risques susceptibles de renforcer la couverture sociale et le développement économique.

Ces évaluations ont démontré que, pris isolément, ces dispositifs entraînaient des impacts significatifs, mais que ceux-ci demeuraient faibles⁶. Dans le cas du microcrédit, cela a contribué à nuancer la portée de cet instrument et à privilégier des approches combinées avec des services d'épargne, de la formation professionnelle et des interventions plus structurelles pour le tissu économique. En ce qui concerne l'assurance santé volontaire, cela a confirmé qu'elle ne constituait un véritable outil de développement et de réduction des inégalités que s'il s'agissait d'un premier pas avant l'élargissement

Enseignements tirés des évaluations d'impact commanditées par l'AFD

Microcrédit Madagascar (2001-2004) : Impact positif en période de croissance de l'économie, mais très incertain en période de récession.

Aménagements des bas-fonds des cultures d'hévéa (1987-2005) : bilan financier négatif pour les producteurs, la technique s'avère accessoire et trop coûteuse.

Téléphonie mobile à Haïti (2008) : Impact mitigé, forte inégalité d'accès, praticité pour des besoins relationnels et l'activité, mais effet d'éviction pour des dépenses primaires.

Microcrédit au Maroc (2006-2009) : Impact limité du microcrédit pris isolément (expansion d'activités existantes). Atteste d'un manque d'adaptation aux besoins des ménages ruraux d'une offre peu diversifiée, centrée sur le crédit.

Prêts étudiants en Afrique du Sud (2011) : lever la contrainte de liquidité augmente significativement les inscriptions et les résultats des étudiants.

Micro-assurance santé au Cambodge (2007-2011) : Impact positif sur la gestion financière et réduction de l'endettement des ménages, mais peu d'éléments probants sur des variables de santé : pointe les limites des modèles d'assurance volontaire.

vers une couverture sociale systématique, ayant vocation à s'étendre à l'ensemble de la population et de susciter des externalités positives sur la qualité des soins.

Au travers de ces exercices, l'AFD a aussi fait progresser la réflexion méthodologique, en pointant le champ d'application restreint des méthodes expérimentales et la difficulté de les combiner avec des approches qualitatives si celles-ci n'avaient pas été intégrées dès le départ comme une composante à part entière. L'AFD a participé à l'évolution du consensus partagé par l'ensemble de la communauté du développement, vers des démarches pluridisciplinaires combinant quantitatif et qualitatif⁷. Elle a en outre montré que les approches les plus rigoureuses ne sauraient constituer un outil systématique de redevabilité, invitant à combler l'écart entre les études d'impact, qui restent rares en raison de leur coût, et les évaluations opérationnelles, qui demeurent souvent restreintes d'un point de vue méthodologique, en développant une gamme diversifiée d'outils de mesure permettant de mieux rendre compte de l'efficacité de l'aide⁸.

Chantiers en cours

A partir de ces enseignements, de nouvelles évaluations d'impact sont en cours de réalisation ou d'instruction. Elles portent sur l'accès à l'eau potable, la santé maternelle et infantile et le soutien aux entreprises.

L'eau, facteur de santé et de bien-être ? L'AFD a commandité deux évaluations dans ce domaine, qui constitue un registre croissant de son intervention. Les incidences de l'accès à l'eau sur la santé (choléra, maladies diarrhéiques, etc.) ou des questions socio-économiques (scolarisation, genre, activité, etc.) ont peu été étudiées sur le continent africain, alors qu'il s'agit d'enjeux étroitement liés aux OMD n° 3, 4 et 5. Des controverses subsistent en outre sur les stratégies les plus efficaces en la matière, par exemple s'il convient pour lutter contre le choléra de privilégier des systèmes d'adduction d'eau potable, de chloration ou de filtration domestique des eaux insalubres, ou de vastes campagnes de vaccination.

Une première étude est actuellement menée sur ces questions par la London School of Hygiene and Tropical Medicine (LSHTM) à Uvira, une ville moyenne du Sud-Kivu (RDC) et mettra l'accent sur les impacts en matière de santé publique. La seconde, encore en cours de montage, portera sur la mise en place d'un système d'adduction d'eau potable faisant l'objet d'une gestion communautaire dans les quartiers périphériques de Kinshasa. Elle s'intéressera en particulier aux impacts socio-économiques et aux effets de ces modes de gestion sur la gouvernance locale.

Ces recherches peuvent être mises en regard avec une autre évaluation d'impact soutenue par l'AFD et menée par DIAL (IRD-Dauphine) à Djibouti et visant à appréhender les effets socioéconomiques d'un projet de rénovation urbaine ciblant un quartier qui concentre la moitié des populations pauvres de la ville. Outre l'accès à l'eau et à l'assainissement, celui-ci prévoit des aménagements relatifs à l'électricité, aux services publics (dispensaire, marché, etc.) et à la formation professionnelle.

Ces évaluations croisent différentes approches, principalement quantitatives. Elles mobilisent à la fois des méthodes expérimentales – aménagement du calendrier des travaux sur la base d'une séquence établie de manière aléatoire, afin de garantir la comparabilité statistique entre les premières zones qui recevront les services (groupe témoin) et les dernières (groupe de contrôle) – et quasi-expérimentales – analyses de données en coupes (données ménages à une échelle plus large que le projet), séries temporelles (ex. historiques d'accès aux services) et appariement pour renforcer la comparabilité entre groupes témoins et contrôles pour que le tirage aléatoire ne soit pas trop contraignant. En cohérence avec les engagements pris à Busan, les modes de collecte de données privilégient en outre le renforcement des dispositifs statistiques locaux (suivi épidémiologique, instituts statistiques, organisme planification urbaine), ce qui a vocation à renforcer leurs capacités à moyen et long terme.

Quelles stratégies pour atteindre les objectifs Muskoka de santé maternelle et infantile ? Une autre évaluation d'impact en cours porte sur le forfait obstétrical mis en œuvre en Mauritanie depuis 2002. Dans un contexte où la plupart des pays sahéliens ont privilégié des politiques de gratuité en matière de santé maternelle, l'approche proposée, qui allie recouvrement des coûts, partage du risque et amélioration de l'offre de soins, est relativement originale. Au démarrage d'une nouvelle extension géographique du projet, il s'agit de mesurer de manière rigoureuse l'impact du forfait en termes d'accès aux soins de santé maternelle ; l'effet du forfait sur les structures sanitaires, en termes de quantité et de qualité des soins, les déterminants et les freins à l'adhésion au forfait et, de façon plus générale, à l'utilisation des services de santé maternelle.

Une démarche évaluative similaire est envisagée en Guinée, où l'AFD collabore avec la Commission européenne pour l'amélioration des services de santé dans la région de Nzérékoré.

De prochaines évaluations menées dans le domaine de la santé seraient susceptibles d'étayer une stratégie sur la planification familiale et la santé reproductive, thèmes qui seront probablement considérés comme prioritaires dans l'agenda post 2015.

Appuis aux entreprises, filières et exportations. Dans le prolongement des travaux menés dans ce domaine sur la microfinance et le développement agricole, l'AFD poursuit une réflexion sur l'évaluation des appuis au secteur privé. Une étude d'impact a ainsi été lancée sur les programmes de mise à niveau des entreprises au Sénégal, reposant sur des données existantes (fichiers de l'Institut National de la Statistique et de la Démographie du Sénégal et données issues du projet) qui seront traitées selon démarche quasi expérimentale pilotée en interne à l'AFD. Les résultats préliminaires sont très positifs au niveau micro (impact sur les entreprises) et les travaux se poursuivent pour établir si des externalités peuvent être décelées au niveau macro (balance commerciale, croissance, etc.)

Par ailleurs, le dispositif ARIZ fait actuellement l'objet d'une évaluation approfondie. On espère qu'elle pourra jeter les bases d'une collecte de données permettant non seulement un meilleur pilotage du dispositif, mais aussi de mener, à terme, une évaluation d'impact permettant d'établir de manière rigoureuse l'additionnalité du dispositif.

3. Bilan et enseignements pour la programmation

L'AFD a dressé en 2014 le bilan de son expérience en matière d'évaluations d'impact⁹, pointant plusieurs vertus de cette activité : elle contribue à la production de connaissances, elle impose un examen rigoureux et approfondi susceptible de favoriser une réflexion critique sur l'efficacité des interventions pour le développement, elle fournit un pendant nuancé aux discours médiatiques qui renforce la crédibilité de l'AFD. Les évaluations d'impact ont par ailleurs des externalités positives pour l'AFD : elles renforcent la notoriété et la légitimité de l'agence dans les forums internationaux, mais également son attractivité pour des ressources humaines ayant un profil scientifique (chercheurs, économistes, spécialistes de l'évaluation, etc.). Elles consolident une culture d'évaluation et suscitent l'attrait des opérationnels à voir rigoureusement évaluer des modes d'interventions sur lesquels ils s'interrogent. Enfin, la mise en place de protocoles rigoureux repose sur un référentiel et des outils pratiques qui permettent d'améliorer la qualité méthodologique des autres évaluations.

Toutefois, malgré l'investissement important qu'elles ont mobilisé, la qualité et l'utilité des résultats obtenus ont été mitigées. Ces évaluations sont utiles pour produire des connaissances scientifiques et stratégiques et pour susciter des apprentissages opérationnels. Elles ne contribuent toutefois que très modestement à la redevabilité de l'AFD : bien qu'elles favorisent la mise en débat et l'appréciation critique de certaines interventions, elles ne permettent de mesurer les résultats que d'interventions simples, qui ne constituent qu'une petite fraction de l'activité de développement de l'agence. En outre, ces exercices requièrent des investissements humains et financiers importants, surtout s'il faut réaliser des enquêtes pour produire des données (entre 300K€ et 800K€ dans ce cas, pour celles financées

jusqu'à présent par l'AFD). Néanmoins, que ce soit à l'AFD, à la BID ou à la Banque mondiale, seule la moitié des initiatives d'évaluations d'impact aboutissent.

B. UNE STRATEGIE A REORGANISER

Dans le prolongement de sa politique d'évaluation, l'AFD souhaite renforcer la pertinence, la fiabilité et l'utilité des évaluations d'impact.

1. Rationaliser l'identification des évaluations d'impact à mener

Il devient indispensable de mieux organiser l'identification, la conduite et l'utilisation des évaluations d'impact à l'AFD. Les premières, initiées par EVA, ont suscité l'intérêt des départements techniques et toutes celles qui sont actuellement en cours ont été lancées à l'initiative des départements techniques. C'est une avancée pour l'appropriation de ces démarches, mais il importe désormais de rationaliser ce processus, afin de renforcer la pertinence, la faisabilité et l'utilité des études. Les services de l'AFD doivent se concerter pour être collectivement en mesure d'identifier les projets qu'il convient d'évaluer.

Critères de sélection des projets à évaluer

Les approches académiques se focalisent souvent sur la démonstration de l'impact (*prove*), laissant de côté l'amélioration des projets ou services (*improve*)¹⁰. Cette priorité se manifeste à différents stades du processus : la définition des questions de recherche, la conception de la méthode, sa mise en œuvre, l'analyse, le format des publications, etc. Favoriser l'exploitation de la recherche pour la prise de décision opérationnelle implique de s'interroger sur les attentes de l'ensemble des parties prenantes du projet et de prendre en compte les différents paramètres qui conditionnent la relation entre le projet avec ses usagers (qualité des services, contexte et antécédents, enjeux socioculturels, etc.). De par sa position de bailleur d'un exercice mettant en jeu des attentes diverses et parfois contradictoires de parties prenantes hétérogènes (décideurs, chercheurs, praticiens, pouvoirs publics locaux, société civile, etc.), l'AFD doit jouer un rôle crucial sur ce registre.

Les évaluations d'impact doivent impérativement être initiées dès l'instruction des projets, notamment parce qu'elles requièrent des données précises sur la situation de référence (*baseline*). Plusieurs critères doivent être pris en compte pour choisir les projets qu'il convient d'évaluer :

Pertinence stratégique : l'intervention évaluée doit revêtir une importance significative pour un cadre d'intervention stratégique ou transversal. Il peut dans cette perspective s'avérer pertinent de lancer simultanément plusieurs évaluations dans un même domaine afin de comparer l'efficacité d'outils alternatifs (ex. garanties vs. ligne de crédit pour les PME) ou d'aborder des volets complémentaires (ex. dimensions sanitaires et socioéconomiques de l'adduction d'eau).

Pertinence scientifique : l'évaluation doit se rapporter à une intervention innovante ou être en rapport avec des questions originales et judicieuses du point de vue de l'état des connaissances ou d'autres recherches en cours.

Intérêt des partenaires : l'évaluation doit susciter l'intérêt de la contrepartie (*a fortiori* lorsqu'il s'agit d'une évaluation financée par le budget du projet), d'autres bailleurs et des parties prenantes locales (autres institutions publiques et privées du pays, société civile, chercheurs et médias locaux), ainsi que des parties prenantes de l'AFD (opérationnels, tutelles, recherche, opinion publique, etc.).

Faisabilité : une phase préparatoire consiste à vérifier la compatibilité entre les temporalités du projet et de l'évaluation, la disponibilité de données, la possibilité de constituer un contrefactuel, les niveaux suffisants d'enrôlement et de maintien des bénéficiaires dans le projet suffisant pour assurer la

significativité des échantillons et la recevabilité éthique de l'exercice. La possibilité d'assurer une gouvernance satisfaisante de l'évaluation (cf. partie 2.3) est une condition fondamentale de leur faisabilité. Cet aspect requiert une attention spécifique dans le cas d'évaluations financées sur projet ou cofinancées par d'autres bailleurs.

Renforcement de capacités : Le recours aux dispositifs de collecte de données existants pour l'évaluation d'impact devrait susciter des externalités positives sur l'intervention évaluée (ex. le système national d'information de santé pour l'évaluation d'une politique), sur les capacités d'évaluations du pays (ex. institut national de statistique) ou sur sa recherche (ex. centres de recherche locaux).

Séquence d'identification

Ces choix dépassent la seule division de l'évaluation et doivent être appréhendés de façon collective. Il convient de ce point de vue de passer d'une approche actuellement opportuniste – on décide de financer ou non des propositions d'évaluation qui nous parviennent –, à un séquençage plus systématique :

- Des champs prioritaires (ex. eau potable, santé maternelle et infantile...) sont définis en concertation entre EVA et la direction des opérations, puis validés par le comité des évaluations, en fonction des enjeux stratégiques pour l'AFD.
- Au fil de l'eau, les opérationnels sont invités à proposer des évaluations lorsqu'un de leurs projets a trait à un champ identifié. Ce critère oriente aussi la veille des CDI par EVA.
- EVA élabore une revue de littérature des études d'impact déjà réalisées dans les champs mis en avant. Cela comporte deux volets :
 - EVA prépare des synthèses sur les sujets déjà couverts ;
 - EVA recense et cartographie et étaye les questions qui restent posées pour les sujets sur lesquels il existe un déficit d'évaluations.
- Les états des lieux des connaissances et des questions en suspens sont présentés aux opérationnels avec des formats adaptés.
- Lorsqu'un projet est retenu, EVA analyse l'opportunité et la faisabilité de son évaluation d'impact. Une mission est souvent requise pour cette phase.
- Un groupe de référence est établi pour orienter et contrôler la progression de l'évaluation. Il assume un rôle de conseil scientifique et sa composition doit être pluraliste (chercheurs, praticiens du développement et décideurs) et pluridisciplinaire.

Une voie alternative : le renforcement du suivi-évaluation

Il peut s'avérer qu'une évaluation d'impact ne rassemble pas les critères de faisabilité ou de pertinence scientifique mentionnés plus haut, mais que l'intervention en question mérite un suivi-évaluation poussé. Si de telles circonstances sont identifiées suffisamment en amont (pendant l'instruction ou au début de la mise en œuvre), EVA peut intervenir en appui au chef de projet et à la maîtrise d'ouvrage pour renforcer le dispositif de collecte, d'analyse et de diffusion d'information.

De tels dispositifs de suivi-évaluation partagent plusieurs aspects des évaluations d'impact : ils sont conçus pour jauger la concrétisation d'une logique d'intervention, ils recueillent des données reflétant les conditions des bénéficiaires à plusieurs stades du projet, l'analyse combine des méthodes quantitatives et qualitatives. Ils sont toutefois plus faciles à mettre en œuvre, car ils ne requièrent pas de contrefactuel. Ils permettent par exemple d'obtenir des renseignements fiables sur le profil socio-économique des publics touchés, leurs besoins qui sont effectivement couverts, leur perception des interventions soutenues par l'AFD et l'évolution de leur niveau de vie, sans toutefois que cette évolution ne puisse être rigoureusement attribuée à l'action de l'AFD.

2. Préciser les modalités essentielles de l'évaluation

Une fois confirmée l'opportunité de l'évaluation, il est indispensable avant son démarrage de fixer les questions de recherche, ainsi que les modalités financières et organisationnelles de l'évaluation, afin que celles-ci répondant aux attentes des parties prenantes, tout en ménageant les conditions indispensables pour mener à bien un tel exercice.

Affiner l'objet de l'évaluation d'impact

Plus le sujet d'une étude d'impact est général, plus celle-ci est coûteuse et moins ses résultats sont clairs, convaincants et utiles. Il faut donc déterminer clairement pour qui et pourquoi on entreprend une étude d'impact afin de préciser les questions évaluatives et d'aboutir à des résultats scientifiquement rigoureux, mais aussi applicables.

Il convient cependant de veiller à bien délimiter le périmètre de l'étude, car l'impact d'interventions complexes ou composites ne peut être évalué avec un seul et même dispositif. Ainsi, l'évaluation d'impact d'un vaste projet alliant formation des paysans, accès aux intrants, aide à la commercialisation et accès au financement requerrait plusieurs études spécifiques, menées simultanément. En outre, les interventions transversales (par exemple le renforcement des capacités du ministère de l'agriculture) se prêtent rarement à une évaluation d'impact au sens propre du terme, car elles recouvrent des appuis hétéroclites mettant en jeu des chaînes de causalité par trop imbriquées et des effets potentiels trop diffus, pour qu'un contrefactuel puisse être isolé. Ce sont donc bien des modes d'interventions spécifiques, et non des projets, qui ont vocation à être évalués. À Kinshasa par exemple, plus que la réussite d'une intervention particulière, l'objectif est de jauger l'efficacité pour le développement local des mécanismes participatifs de gestion de l'eau en milieu péri-urbain. Cette portée des évaluation au-delà d'un projet particulier est indispensable pour justifier une approche scientifiquement poussée et coûteuse.

Assurer le financement et la gouvernance des évaluations

Il est indispensable qu'un acteur ait un rôle proactif de pilotage de ces études, semblable à une maîtrise d'ouvrage, afin d'assurer la coordination entre chercheurs et opérateurs. En effet, la recherche d'un contrefactuel valide peut impliquer de nombreuses contraintes pour le projet : maîtrise du calendrier des opérations, formalisation et aménagement du mode de sélection des bénéficiaires, etc. Une entité doit être en mesure d'arbitrer les tensions éventuelles pour trouver des solutions constructives. Sinon les chercheurs feront leur étude en fonction de leur agenda au détriment des opérations et des attentes des parties prenantes (ex. étude sur l'assurance santé au Cambodge) ou les opérateurs privilégieront leurs propres priorités, au détriment de l'évaluation. Ce hiatus est l'une des principales causes d'échec de ce type d'études.

En raison des contraintes budgétaires de l'AFD, les enquêtes de terrain – qui représentent le principal coût des évaluations d'impact – sont désormais presque systématiquement financées sur les budgets des projets. Ce mode de financement tend *a priori* à accorder un rôle de maîtrise d'ouvrage aux contreparties locales, ce qui fait peser un risque sur la réussite des évaluations lorsque les administrations ne disposent pas des compétences et des procédures ou des motivations requises.

Pour cette raison, la division évaluation de l'AFD plaide pour qu'un rôle d'assistance à la maîtrise d'ouvrage lui soit dévolu pour les évaluations d'impact. Dans ce cadre, un accord formel et explicite doit clarifier les modes de gouvernance de l'évaluation et sa signature conditionne la faisabilité de l'étude du point de vue d'EVA. Lorsque les risques liés aux capacités de la maîtrise d'ouvrage sont jugés excessifs, une délégation de la maîtrise d'ouvrage à EVA doit être demandée, comme cela a été fait pour certains projets menés auprès d'États fragiles.

Clarifier la gouvernance des études implique aussi de s'entendre sur le mode d'attribution des marchés relatifs à l'évaluation. En effet, bien qu'une mise en concurrence soit souhaitable, la passation d'appels d'offres « classiques » n'est souvent pas adaptée à ce type d'exercice, car elle attire essentiellement des consultants et bureaux d'études. Or, le constat partagé par plusieurs commanditaires d'études d'impact est qu'il est préférable de recourir à des chercheurs universitaires pour avoir des résultats solides tant au niveau quantitatif que qualitatif. Ceux-ci répondent cependant plus rarement à des appels d'offres. En outre, les procédures classiques de mise en concurrence semblent d'autant moins adaptées qu'un travail considérable de préparation est nécessaire pour une évaluation d'impact et que le contenu de l'évaluation (et donc son coût) est susceptible d'être substantiellement ajusté après des travaux préliminaires. Pour toutes ces raisons, l'AFD prévoit dans sa politique d'évaluation la possibilité de mener ce type d'études grâce à des partenariats de recherches reposant en priorité sur des accords de gré à gré et non sur des appels d'offres.

Il importe enfin de préciser dès le démarrage la question de la propriété intellectuelle et de la diffusion des produits de l'évaluation, qui doivent constituer des biens publics. Cela renvoie tout d'abord aux résultats finaux, qui doivent être aussi diffusés sous des formats accessibles aux opérationnels, aux décideurs et aux publics, dans les pays récipiendaires et parmi les bailleurs. Cela concerne également les produits intermédiaires (questionnaires, outils d'enquête, etc.), qui doivent être rendus accessibles à une échéance fixée dès le départ.

C. DES APPROCHES COMPLEMENTAIRES

Parallèlement aux efforts visant à optimiser la pertinence et l'utilité des évaluations d'impact menées par l'AFD, trois fonctions complémentaires peuvent être développées.

1. Un rôle de courtier de connaissances

L'atteinte des objectifs de production d'enseignements utiles à la stratégie et aux opérations de l'AFD implique que la division de l'évaluation acquière une fonction de « courtier de connaissances ». Ces efforts permettront de mieux cadrer de futurs projets d'évaluation d'impact et de s'assurer que les enseignements des évaluations disponibles soient effectivement diffusés auprès des opérationnels et décideurs de l'AFD.

EVA doit ainsi assurer une fonction de collecte, de synthèse et de diffusion de connaissances sur l'impact des interventions de développements à partir d'évaluation menées par d'autres. Des livrables tels que des revues de littératures, des méta-évaluations, et des animations scientifiques, pourront ainsi fournir aux divisions opérationnelles et stratégiques des éléments susceptibles de guider leurs choix, de mieux concevoir leurs projets et d'en valoriser les résultats.

2. L'animation de partenariats scientifiques

Outre les études commanditées par l'AFD, il semble aussi utile qu'EVA incite d'autres acteurs impliqués dans l'organisation ou la conduite d'évaluation d'impact à orienter leurs activités vers les thématiques et les géographies d'intérêt de l'agence. Cela concerne en particulier les équipes de chercheurs avec lesquels des partenariats institutionnels ont été tissés (CEPREMAP, IRD, Science Po, CERDI, etc.) et qui pourraient être encouragés à s'intéresser aux projets de l'AFD pour en évaluer l'impact, grâce à un accès – facilité par EVA – aux terrains et aux données. EVA jouerait dans ce cadre un rôle de médiateur pour la définition de questions évaluatives, la conception de protocoles, la collecte de données, ou encore la diffusion des résultats. Ce type de partenariats pourrait à terme déboucher sur le montage de projets communs présentés à des guichets de financement de la recherche français (Agence Nationale de la Recherche) et européens (FP7).

3. La contribution aux débats méthodologiques

Cet effort doit porter sur les méthodes, afin de faire progresser les approches mixtes (conjuguant quantitatif et qualitatif), mais aussi pour renforcer et mieux exploiter les données statistiques nationales ou les systèmes d'informations locaux.

Les enjeux de méthode requièrent une attention particulièrement poussée dans le domaine des évaluations d'impact, car elles déterminent leur faisabilité, leur coût, leur recevabilité et leur portée. Il semble donc indispensable que l'AFD consolide ses propres capacités dans ce sens, et qu'elle continue à se positionner l'AFD dans le débat d'idées sur la production de connaissances sur le développement.

Le champ des évaluations d'impact est caractérisé depuis une décennie par la montée en puissance des études expérimentales. Bien qu'il se fonde sur plusieurs avancées méthodologiques certaines, cet engouement masque les limites des expérimentations : application restreinte aux programmes « tunnels »¹¹, distorsions et jeux d'acteurs qui peuvent biaiser leurs données¹². Tout en reconnaissant l'intérêt des expérimentations, il faut pointer les limites des évaluations s'appuyant exclusivement sur l'assignation aléatoire et défendre des approches mixtes combinant méthodes quantitatives et qualitatives.

Cette articulation est invoquée de manière presque rituelle, mais elle est rarement mise en œuvre, pour des raisons aussi bien institutionnelles (revues, évaluation et de sources financement de la recherche), que techniques (procédés de collecte d'information et compétences requises pour l'analyse), épistémologiques (faits objectifs/perceptions) et dans leurs modes d'administration de la preuve (significativité statistique/plausibilité et cohérence, démarche hypothético-déductive/inductive). Ces chercheurs évoluent donc dans des référentiels différents, et ont du mal à s'accorder sur des questions de recherche, des processus d'investigation et des formats de publications qui permettraient réellement d'orienter l'amélioration des interventions. De par sa position, l'AFD a vocation à jouer un rôle de pionnier dans le cadre des évaluations qu'elle organise, contribuant à faire progresser les méthodes d'évaluation d'impact des interventions de développement.

CONCLUSION ET PROCHAINES ETAPES

Les questions posées par Direction générale et du Comité des évaluations à la division EVA l'ont conduit à actualiser sa stratégie d'évaluations d'impact. Un travail d'inventaire et des consultations auprès des acteurs concernés au sein de l'agence ont permis de tirer les enseignements d'une décennie d'études dans ce domaine. Ces apprentissages incitent à se fixer des objectifs modestes quant au nombre d'évaluations d'impact commanditées par l'agence, mais de consacrer les moyens nécessaires pour que ces études soient de qualité. Quatre critères doivent être réunis pour qu'une évaluation d'impacts commandités par l'AFD se déroule dans de bonnes conditions :

1. Que les évaluations soient conçues dès l'instruction des projets au travers d'une procédure claire et sur la base de critères partagés DOE-EVA ;
2. Qu'un financement suffisant soit disponible pour l'ensemble de l'évaluation ;
3. Que la maîtrise d'ouvrage revienne – directement ou par délégation – à EVA ;
4. Que soit possible le recours au gré à gré avec des laboratoires publics et/ou des instituts nationaux de statistique, au travers de partenariats de recherche.

Ces conditions sont difficilement remplies lorsque le financement de l'évaluation est prévu sur le budget du projet à évaluer. Le Comité des évaluations recommande sur ce point de viser un objectif raisonnable équivalent à une évaluation d'impact tous les deux ans, mais de ce donner les moyens de les financer sur le budget de l'AFD. Cela impliquerait toutefois de doubler le montant actuellement consacré aux évaluations d'impact pour atteindre environ 200 K€/an.

Cette activité doit s'accompagner d'un rôle actif en matière de gestion des connaissances, d'animation de partenariats de recherche et d'implication dans les débats internationaux. La prise en compte des controverses méthodologiques et des enjeux scientifiques ne doivent cependant pas conduire EVA à délaisser les préoccupations d'apprentissages sur les pratiques et les stratégies de l'AFD, qui constituent l'objectif prioritaire poursuivi au travers des évaluations d'impact. Les quatre piliers mis en avant dans la stratégie d'évaluation d'impact doivent donc s'articuler pour sous-tendre un agenda de recherche cohérent et dynamique, élaboré conjointement avec les opérationnels.

Les jalons pour la suite du processus sont les suivants :

- Identifier avec la DOE un nombre restreint de questions de développement cristallisant à la fois des enjeux scientifiques pour la communauté des chercheurs et décideurs sur le développement et des enjeux stratégiques pour les opérationnels de l'AFD (ex. gratuité des services de base, planification familiale, modes de gestion de l'éducation...);
- Mener des revues systématiques des évaluations d'impact disponibles dans ces domaines, en se focalisant sur des aires géographiques où l'AFD intervient ;
- Concevoir et mener une évaluation sur le thème qui semble le plus pertinent, en mobilisant une approche alliant quantitatif et qualitatif ;
- Capitaliser sur cette expérience pour peser dans les débats internationaux en matière de recherche évaluative et de résultats de développement.

REFERENCES

¹ Pamies Sumner Stéphanie, *Les évaluations d'impact dans le domaine du développement : état des lieux et nouveaux enjeux*, Paris, Agence Française de Développement, À Savoir, juin 2014.

² Comité d'aide au développement, *Glossaire des principaux termes relatifs à l'évaluation et la gestion axée sur les résultats*, Paris, OCDE, 2002, p.24, disponible à l'adresse : <http://www.afd.fr/webdav/shared/RECHERCHE/Evaluation/document/glossaire.pdf>

³ Voir sur le site de l'AFD la page *Outils de mesure des résultats et impacts*, disponible à l'adresse : <http://www.afd.fr/home/AFD/developpement-durable/Mesure-par-les-resultats/outils-mesure-impacts>

⁴ Jean-David Naudet, Jocelyne Delarue, *Fostering Impact Evaluations at Agence Française de Développement: A Process of In-house Appropriation and Capacity-Building*, Washington DC, Network of Networks on Impact Evaluation, Working Paper, janv. 2008.

⁵ Elizabeth Littlefield, Jonathan Morduch, Syed Hashemi, *Is microfinance an effective strategy to reach the Millenium Development Goals ?*, vol. (24), Washington DC, CGAP, janv. 2003.

⁶ Bruno Crépon et al., *Impact of microcredit in rural areas of Morocco: Evidence from a Randomized Evaluation*, Paris, CREST, Working Paper, mars 2011 ; Rachel Polimeni, David I. Levine, *Going Beyond Adverse Selection: Take-up of a Health Insurance Program in Rural Cambodia*, Paris, AFD, Ex Post-Analyses d'impact 11, 2013.

⁷ Jean-David Naudet, Jocelyne Delarue, Tanguy Bernard, « Évaluations d'impact : un outil de redevabilité ? Les leçons tirées de l'expérience de l'AFD », *Revue d'économie du développement*, Vol. 26 (4), 2012, pp. 27-48 ; Stéphanie Pamies Sumner, *Les évaluations d'impact dans le domaine du développement : quelle stratégie pour un bailleur comme l'AFD ?*, Paris, AFD, Version provisoire présentée au Comité des Évaluations, oct. 2013.

⁸ Agence française de développement (Éd.), *La politique d'évaluation de l'AFD*, Paris, AFD, 2013.

⁹ Pamies Sumner, *Les évaluations d'impact dans le domaine du développement*, op. cit.

¹⁰ David Hulme, « Impact Assessment Methodologies for Microfinance: Theory, Experience and Better Practice », *World Development*,, 28 (1), 2000, pp. 79-98.

¹¹ Naudet, Delarue, Bernard, « Évaluations d'impact », *op. cit.*

¹² Aurélie Quentin, Isabelle Guérin, « La randomisation à l'épreuve du terrain », *Revue Tiers Monde*, n°213 (1), mars 2013, pp. 179-200.